

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:41:17 ; Search time 23 Seconds
(without alignments)
2558.852 Million cell updates/sec

Title: US-10-045-072-2

Perfect score: 5788

Sequence: 1 MSTHTSSTLPFAFKKILVANR.....RVVVPAATKVEGGDLIVVVS 1140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	5788	100.0	1140	3	US-09-220-081-2
2	5788	100.0	1140	4	US-09-677-575-2
3	3483	60.2	1124	4	US-08-311-731A-10
4	2464	42.6	1154	4	US-09-134-001C-3428
5	2463.5	42.6	1163	4	US-09-134-000C-5707
6	2174	37.6	973	4	US-09-107-532A-4810
7	1192	20.6	694	4	US-09-433-043B-126
8	1083	18.7	456	4	US-09-634-238-276
9	1043.5	18.0	593	4	US-09-433-043B-122
10	1036.5	17.9	447	1	US-08-611-107-6
11	1036.5	17.9	447	2	US-08-422-560A-6
12	1036.5	17.9	447	3	US-08-468-793-6
13	1004.5	17.4	453	1	US-08-611-107-8
14	1004.5	17.4	453	2	US-08-422-560A-8
15	1004.5	17.4	453	3	US-08-468-793-8
16	1002.5	17.3	453	4	US-09-433-043B-121
17	995.5	17.2	453	1	US-07-956-700B-6
18	995.5	17.2	453	1	US-08-476-537-6
19	995.5	17.2	453	1	US-08-485-607-6
20	995.5	17.2	453	2	US-08-475-879-6
21	995.5	17.2	453	4	US-09-433-043B-6
22	967.5	16.7	474	4	US-09-328-352-7562
23	966.5	16.7	1116	4	US-09-252-991A-24374
24	961	16.6	453	4	US-09-543-681A-5871
25	952	16.4	605	4	US-09-433-043B-123
26	944.5	16.3	451	4	US-09-540-236-3047
27	938	16.2	448	1	US-08-074-121-3

28	938	16.2	448	5	PCT-US94-06447-3	Sequence 3, Appli
29	937	16.2	427	1	US-07-956-700B-3	Sequence 3, Appli
30	937	16.2	427	1	US-08-476-537-3	Sequence 3, Appli
31	937	16.2	427	1	US-08-485-607-3	Sequence 3, Appli
32	937	16.2	427	2	US-08-475-879-3	Sequence 3, Appli
33	937	16.2	427	4	US-09-433-043B-3	Sequence 3, Appli
34	935.5	16.2	454	4	US-09-198-452A-197	Sequence 197, App
35	935.5	16.2	701	4	US-09-252-991A-27999	Sequence 27999, A
36	926	16.0	536	3	US-08-662-344-2	Sequence 2, Appli
37	924.5	16.0	465	4	US-09-252-991A-26980	Sequence 26980, A
38	924	16.0	449	1	US-08-074-121-6	Sequence 6, Appli
39	924	16.0	449	5	PCT-US94-06447-6	Sequence 6, Appli
40	914	15.8	457	4	US-09-134-000C-4554	Sequence 4554, Ap
41	901	15.6	676	4	US-09-252-991A-26143	Sequence 26143, A
42	899	15.5	670	4	US-09-328-352-6725	Sequence 6725, Ap
43	896	15.5	619	4	US-09-543-681A-5852	Sequence 5852, Ap
44	891	15.4	652	4	US-09-328-352-5587	Sequence 5587, Ap
45	889	15.4	573	4	US-09-328-352-6420	Sequence 6420, Ap

ALIGNMENTS

RESULT 1
US-09-220-081-2
; Sequence 2, Application US/09220081
; Patent No. 6171833
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; APPLICANT: Stephanopoulos, Gregory
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790000
; CURRENT APPLICATION NUMBER: US/09/220,081
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-220-081-2

Query Match	100.0%;	Score 5788;	DB 3;	Length 1140;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1140;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSTHTSSTLPFAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR	60	
Db	1	MSTHTSSTLPFAFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR	60	
QY	61	IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV	120	
Db	61	IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV	120	
QY	121	LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGOTYPIFVKAVAGGGGRGMR	180	
Db	121	LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGOTYPIFVKAVAGGGGRGMR	180	
QY	181	VASPELRLKLAATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVHLYERDCS	240	
Db	181	VASPELRLKLAATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVHLYERDCS	240	
QY	241	LQRRHQKVVEIAPAQHLDPQLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEVN	300	
Db	241	LQRRHQKVVEIAPAQHLDPQLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEVN	300	
QY	301	PRIQVEHTVTEEVTEVDLVKAQVRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG	360	
Db	301	PRIQVEHTVTEEVTEVDLVKAQVRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG	360	
QY	361	FRPDTGITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL	420	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:44:22 ; Search time 50 Seconds
(without alignments)
S904.182 Million cell updates/sec

Title: US-10-045-072-2
Perfect score: 5788
Sequence: 1 MSTHTSSTLPAPFKILVANR.....RVVPAATKVEGGDLIVVVS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5788	100.0	1140	9	US-09-974-973-19 Sequence 19, Appl
2	5788	100.0	1140	9	US-09-738-626-4265 Sequence 4265, Ap
3	5788	100.0	1140	14	US-10-045-072-2 Sequence 2, Appli
4	5759	99.5	1157	9	US-09-974-973-2 Sequence 2, Appli
5	5759	99.5	1157	9	US-09-974-973-4 Sequence 4, Appli
6	4647.5	80.3	1141	12	US-10-282-122A-53885 Sequence 53885, A
7	3791.5	65.5	1127	12	US-10-282-122A-61950 Sequence 61950, A
8	3712.5	64.1	1127	12	US-10-282-122A-62684 Sequence 62684, A
9	3712.5	64.1	1127	12	US-10-282-122A-64810 Sequence 64810, A
10	2552	44.1	1144	12	US-10-282-122A-52777 Sequence 52777, A
11	2526	43.6	1148	12	US-10-282-122A-46866 Sequence 46866, A
12	2525.5	43.6	1150	15	US-10-369-493-17338 Sequence 17338, A
13	2524	43.6	1148	15	US-10-369-493-23112 Sequence 23112, A
14	2493	43.1	1146	12	US-10-282-122A-60488 Sequence 60488, A
15	2491.5	43.0	1144	12	US-10-282-122A-52088 Sequence 52088, A

16	2491.5	43.0	1144	15	US-10-369-493-7766 Sequence 7766, Ap
17	2488	43.0	1147	9	US-09-815-242-5468 Sequence 5468, Ap
18	2488	43.0	1150	12	US-10-282-122A-44391 Sequence 44391, A
19	2472.5	42.7	1142	9	US-09-815-242-10806 Sequence 10806, A
20	2471.5	42.7	1147	15	US-10-369-493-11450 Sequence 11450, A
21	2471.5	42.7	1151	15	US-10-369-493-14817 Sequence 14817, A
22	2471.5	42.7	1151	15	US-10-369-493-15006 Sequence 15006, A
23	2470.5	42.7	1152	15	US-10-369-493-12027 Sequence 12027, A
24	2464.5	42.6	1142	12	US-10-282-122A-57942 Sequence 57942, A
25	2464	42.6	1175	15	US-10-369-493-6504 Sequence 6504, Ap
26	2463.5	42.6	1142	12	US-10-282-122A-42528 Sequence 42528, A
27	2457	42.4	1178	15	US-10-369-493-21939 Sequence 21939, A
28	2443	42.2	1167	15	US-10-369-493-22819 Sequence 22819, A
29	2443	42.2	1185	15	US-10-369-493-2488 Sequence 2488, Ap
30	2436	42.1	1180	15	US-10-369-493-1491 Sequence 1491, Ap
31	2406	41.6	1137	15	US-10-369-493-18351 Sequence 18351, A
32	2375	41.0	1143	12	US-10-282-122A-52917 Sequence 52917, A
33	2299	39.7	1073	9	US-09-815-242-12361 Sequence 12361, A
34	2185.5	37.8	1385	15	US-10-369-493-3805 Sequence 3805, Ap
35	2096.5	36.2	903	15	US-10-369-493-14300 Sequence 14300, A
36	1827.5	31.6	1076	15	US-10-369-493-13831 Sequence 13831, A
37	1377.5	23.8	1122	12	US-10-282-122A-46406 Sequence 46406, A
38	1297	22.4	641	12	US-10-282-122A-57483 Sequence 57483, A
39	1290.5	22.3	611	12	US-10-282-122A-70702 Sequence 70702, A
40	1285.5	22.2	611	12	US-10-282-122A-70190 Sequence 70190, A
41	1276.5	22.1	611	12	US-10-282-122A-71785 Sequence 71785, A
42	1245.5	21.5	596	15	US-10-369-493-8430 Sequence 8430, Ap
43	1212.5	20.9	716	15	US-10-369-493-12942 Sequence 12942, A
44	1210.5	20.9	581	15	US-10-369-493-19178 Sequence 19178, A
45	1156	20.0	454	15	US-10-369-493-13206 Sequence 13206, A

ALIGNMENTS

RESULT 1
US-09-974-973-19
; Sequence 19, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteri
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; PRIOR FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-19

Query Match	100.0%;	Score 5788;	DB 9;	Length 1140;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1140;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSTHTSSTLPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVR	60	
Db	1	MSTHTSSTLPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVR	60	
QY	61	IGTEGSPVKAYLDIDEIIGAACKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEV	120	
Db	61	IGTEGSPVKAYLDIDEIIGAACKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEV	120	
QY	121	LDLTGDKSRVATAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF	180	
Db	121	LDLTGDKSRVATAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF	180	
QY	181	VASPELRLKLTAEASREAEAFGDGAVYVERAVINPOHIEVQILGHTGEVWHLYERDCS	240	
Db	181	VASPELRLKLTAEASREAEAFGDGAVYVERAVINPOHIEVQILGHTGEVWHLYERDCS	240	

Db 181 VASDELRLKLATEASREAEAAFGDGA VYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
QY 241 LQRRHQKVVEIAPAQHLDPRLDRICADAVKFCRSIGYQGAGTVFVLVDEKGNHVFIEMN 300
Db 241 LQRRHQKVVEIAPAQHLDPRLDRICADAVKFCRSIGYQGAGTVFVLVDEKGNHVFIEMN 300
QY 301 PRIQVEHTVTEEVTEVDLVKQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
Db 301 PRIQVEHTVTEEVTEVDLVKQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
QY 361 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db 361 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
QY 541 FRDAHQSILLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWRLDE 600
Db 541 FRDAHQSILLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWRLDE 600
QY 601 LREAMPNVNIQMLLRNRTVGTYPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRNRTVGTYPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
QY 661 IDAVLETNTAVAEVAMAYSGDSDPNNEKLYTLDYLYLKMAEEIVKSGAHILAIDKDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGDSDPNNEKLYTLDYLYLKMAEEIVKSGAHILAIDKDMAGLLR 720
QY 721 PAAVTKLVTALRREFDLPVHVHTHTAGGOLATYFAAAQAGADAVDGASAPLSGTTSPQS 780
Db 721 PAAVTKLVTALRREFDLPVHVHTHTAGGOLATYFAAAQAGADAVDGASAPLSGTTSPQS 780
QY 781 LSAIIVAFAFHTRRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db 781 LSAIIVAFAFHTRRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPESSKVVGDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPESSKVVGDLALHLVGAGVDPADF 900
QY 901 AADPOKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHLDA 960
Db 901 AADPOKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHLDA 960
QY 961 DDSKERNSLNRLLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
Db 961 DDSKERNSLNRLLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
QY 1021 LLVRLDAISEPDDKGMNRNVVANVNGQIRPMRVDRDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db 1021 LLVRLDAISEPDDKGMNRNVVANVNGQIRPMRVDRDRSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1081 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140

RESULT 2
US-09-738-626-4265
; Sequence 4265, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4265
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4265

Query Match 100.0%; Score 5788; DB 9; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTHTSSTLPFAFKKILVANRGEIIVRAFRALLETGAATVAIYPREDRGSFHRSEAVR 60
Db 1 MSTHTSSTLPFAFKKILVANRGEIIVRAFRALLETGAATVAIYPREDRGSFHRSEAVR 60
QY 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGVCFLSENAQLARECAENGITFIGPTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGVCFLSENAQLARECAENGITFIGPTPEV 120
QY 121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180
Db 121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180
QY 181 VASDELRLKLATEASREAEAAFGDGA VYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
Db 181 VASDELRLKLATEASREAEAAFGDGA VYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
QY 241 LQRRHQKVVEIAPAQHLDPRLDRICADAVKFCRSIGYQGAGTVFVLVDEKGNHVFIEMN 300
Db 241 LQRRHQKVVEIAPAQHLDPRLDRICADAVKFCRSIGYQGAGTVFVLVDEKGNHVFIEMN 300
QY 301 PRIQVEHTVTEEVTEVDLVKQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
Db 301 PRIQVEHTVTEEVTEVDLVKQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNN 360
QY 361 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db 361 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
QY 541 FRDAHQSILLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWRLDE 600
Db 541 FRDAHQSILLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWRLDE 600
QY 601 LREAMPNVNIQMLLRNRTVGTYPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRNRTVGTYPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
QY 661 IDAVLETNTAVAEVAMAYSGDSDPNNEKLYTLDYLYLKMAEEIVKSGAHILAIDKDMAGLLR 720

Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTLRREFDLPVHVHTHTTAGGQLATYFAAAQAGADAVD GASAPLSGTTSQPS 780
Db 721 PAAVTKLVTLRREFDLPVHVHTHTTAGGQLATYFAAAQAGADAVD GASAPLSGTTSQPS 780
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESTPGPTGRVYRHEIPGGQL 840
Db 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESTPGPTGRVYRHEIPGGQL 840
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
QY 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPL TEVPEEEQAH LDA 960
Db 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPL TEVPEEEQAH LDA 960
QY 961 DSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db 961 DSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
QY 1021 LLVRLDAISEPDDKGMNVNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAA PFA 1080
Db 1021 LLVRLDAISEPDDKGMNVNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAA PFA 1080
QY 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAA TKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAA TKVEGGDLIVVVS 1140

RESULT 3

US-10-045-072-2
; Sequence 2, Application US/10045072
; Publication No. US20030027305A1
; GENERAL INFORMATION:
; APPLICANT: Sinskey, Anthony J.
; APPLICANT: Lessard, Philip A.
; APPLICANT: Willis, Laura B.
; TITLE OF INVENTION: Pyruvate Carboxylase from Corynebacterium glutamicum
; FILE REFERENCE: 1533.0790002
; CURRENT APPLICATION NUMBER: US/10/045,072
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 09/677,575
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US 09/220,081
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-045-072-2

Query Match 100.0%; Score 5788; DB 14; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAPKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
Db 1 MSTHTSSTLPAPKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
QY 61 ICTEGSPVKAYLDIDEIIGAARKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEV 120
Db 61 ICTEGSPVKAYLDIDEIIGAARKVKADAIYPGVGFLSENAQLARECAENGITFIGPTPEV 120
QY 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMR 180
Db 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMR 180
QY 181 VASPDLEKRLATEASREAEAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240

Db 181 VASPDLEKRLATEASREAEAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
QY 241 LQRRHOKVVEIAPAQHLDPQLRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIE MN 300
Db 241 LQRRHOKVVEIAPAQHLDPQLRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIE MN 300
QY 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
QY 361 FRPDTGTTTAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQAL 420
Db 361 FRPDTGTTTAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDPLPRGSRDLKQLGPAAFAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAPIDKLPNIKDPLPRGSRDLKQLGPAAFAFARDLREQDALAVTDTT 540
QY 541 FRDAHQSLLATRVRSFALKPAAEAFAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDR LDE 600
Db 541 FRDAHQSLLATRVRSFALKPAAEAFAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDR LDE 600
QY 601 LREAMPNVNIQMLLRGNTVGYTPYDSCVAFVKEAASSGVDFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRGNTVGYTPYDSCVAFVKEAASSGVDFRIFDALNDVSQMRPA 660
QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTLRREFDLPVHVHTHTTAGGQLATYFAAAQAGADAVD GASAPLSGTTSQPS 780
Db 721 PAAVTKLVTLRREFDLPVHVHTHTTAGGQLATYFAAAQAGADAVD GASAPLSGTTSQPS 780
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESTPGPTGRVYRHEIPGGQL 840
Db 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESTPGPTGRVYRHEIPGGQL 840
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
QY 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPL TEVPEEEQAH LDA 960
Db 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPL TEVPEEEQAH LDA 960
QY 961 DSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db 961 DSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
QY 1021 LLVRLDAISEPDDKGMNVNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAA PFA 1080
Db 1021 LLVRLDAISEPDDKGMNVNVANVNGQIRPMVRDRSVESVTATAEKADSSNKGHVAA PFA 1080
QY 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAA TKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAA TKVEGGDLIVVVS 1140

RESULT 4

US-09-974-973-2
; Sequence 2, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteri
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; CURRENT FILING DATE: 2001-10-21


```
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-2

Query Match      99.5%; Score 5759; DB 9; Length 1157;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSPASEAVR 60
Db :
18 VSTHTSSTLPFAFKKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSPASEAVR 77

61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
Db :
78 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137

121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMR 180
Db :
138 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMR 197

181 VASPDRLKRLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240
Db :
198 VSSPDRLKRLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCS 257

241 LQRRHQKVEIAPAQHLDPBLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300
Db :
258 LQRRHQKVEIAPAQHLDPBLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 317

301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLTKELGLTQDKIKTHGAALQCRITTEDPNN 360
Db :
318 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLTKELGLTQDKIKTHGAALQCRITTEDPNN 377

361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db :
378 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437

421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
Db :
438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 497

481 TVNKPVGVRPKDVAAPIDKLPNIDPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
Db :
498 TVNKPVGVRPKDVAAPIDKLPNIDPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 557

541 FRDAHQSLLATRVRSFALKPAABAVAKLTPELLSVEAWGATYDVAMRFLFEDPDWRLLDE 600
Db :
558 FRDAHQSLLATRVRSFALKPAABAVAKLTPELLSVEAWGATYDVAMRFLFEDPDWRLLDE 617

601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFVKEAASSGVDIFRIFDALNDVSMRPA 660
Db :
618 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFVKEAASSGVDIFRIFDALNDVSMRPA 677

661 IDAVLETTNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIDMAGLLR 720
Db :
678 IDAVLETTNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIDMAGLLR 737

721 PAAVTKLVTLARREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTSQPS 780
Db :
738 PAAVTKLVTLARREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTSQPS 797

781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db :
798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 857

841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDALHLVGAGVDPADF 900
Db :
858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDALHLVGAGVDPADF 917
```

```
QY 901 AADPQKYDIPDSVIAFLRGLGNPPGPGWPEPLRTRALEGRSEKAPLTEVPEEEQAHLDA 960
Db :
918 AADPQKYDIPDSVIAFLRGLGNPPGPGWPEPLRTRALEGRSEKAPLTEVPEEEQAHLDA 977

961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
Db :
978 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1037

1021 LLVRLDAISEPDDKGMNRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db :
1038 LLVRLDAISEPDDKGMNRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1097

1081 GWVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db :
1098 GWVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERVVVPAATKVEGGDLIVVVS 1157

RESULT 5
US-09-974-973-4
; Sequence 4, Application US/09974973
; Patent No. US20020177202A1
; GENERAL INFORMATION:
; APPLICANT: Hanke, Paul D.
; TITLE OF INVENTION: Feedback-Resistant Pyruvate Carboxylase Gene from Corynebacteriu
; FILE REFERENCE: 1533.1230001/MAC/RGM
; CURRENT APPLICATION NUMBER: US/09/974,973
; PRIOR FILING DATE: 2001-10-21
; PRIOR APPLICATION NUMBER: US 60/239,913
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1157
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-974-973-4

Query Match      99.5%; Score 5759; DB 9; Length 1157;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSPASEAVR 60
Db :
18 VSTHTSSTLPFAFKKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSPASEAVR 77

61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
Db :
78 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 137

121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMR 180
Db :
138 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDDIVKSAEGQTYPIFVKAVAGGGGRGMR 197

181 VASPDRLKRLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240
Db :
198 VSSPDRLKRLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTGEVHLYERDCS 257

241 LQRRHQKVEIAPAQHLDPBLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300
Db :
258 LQRRHQKVEIAPAQHLDPBLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 317

301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLTKELGLTQDKIKTHGAALQCRITTEDPNN 360
Db :
318 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLTKELGLTQDKIKTHGAALQCRITTEDPNN 377

361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db :
378 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 437

421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
Db :
438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 497
```

QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVDTT 540
Db TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVDTT 557
QY 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 600
Db FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 617
QY 601 LREAMPNVIQMLLRGRNTVGYTPYDPSVCRAFVKEAAASSGVDFRIFDALNDVSQMRPA 660
Db LREAMPNVIQMLLRGRNTVGYTPYDPSVCRAFVKEAAASSGVDFRIFDALNDVSQMRPA 677
QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEELVKSGAHILAIKDMAGLLR 720
Db IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEELVKSGAHILAIKDMAGLLR 737
QY 721 PAAVTKLVTLRRREFDLPVHVHTHTAGGQLATYFAAAQACADAVDGASAPLSGTSQPS 780
Db PAAVTKLVTLRRREFDLPVHVHTHTAGGQLATYFAAAQACADAVDGASAPLSGTSQPS 797
QY 781 LSAIVAFAFHTRRDTGLSLEAVSDLEPYWEAVRGLYLPPEFSGTPGPTGRVYRHEIPGGQL 840
Db LSAIVAFAFHTRRDTGLSLEAVSDLEPYWEAVRGLYLPPEFSGTPGPTGRVYRHEIPGGQL 857
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTSPSKVVGDLALHLVGAGVDPADP 900
Db SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTSPSKVVGDLALHLVGAGVDPADP 917
QY 901 AADPQKYDIPDSVIAFLRGEIPLNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHILDA 960
Db AADPQKYDIPDSVIAFLRGEIPLNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHILDA 977
QY 961 DDSKERRNSLNRLLFPKPTTEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db DDSKERRNSLNRLLFPKPTTEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1037
QY 1021 LLVRLDAISEPDDKGMNVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db LLVRLDAISEPDDKGMNVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1097
QY 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
Db GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIERVVVVPAATKVEGGDLIVVVS 1157

RESULT 6
US-10-282-122A-53885
; Sequence 53885, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 53885
; LENGTH: 1141
; TYPE: PR1
; ORGANISM: *Corynebacterium diptheriae*
US-10-282-122A-53885

Query Match 80.3%; Score 4647.5; DB 12; Length 1141;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 917; Conservative 81; Mismatches 137; Indels 3; Gaps 2;

QY 2 STHTSSTLPAPKKILVANRGEIAVRAAFRAALETGAATVAIYPREDRSGFHRSEASEAVRI 61
Db 6 SERGTSTVNPLSKILVANRGEIAVRAAFRAAFETGAATVAVPNEDRNSFHRSEASEAVLI 65
QY 62 GTEGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVL 121
Db 66 GEGGSVVKAYLDIDEIIRAQKQTGADAIYPGYGFLSENAQLARECAENGITFIGPPPSVL 125
QY 122 DLTGDKSRVATAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFV 181
Db 126 ELTGDKAAAVTAAREAGLPTLTETEATDDPKKLAELTKGQTYPLFVKAVAGGGGRGMRFV 185
QY 182 ASPDELRLKLAATESREAAFGDGVYVERAVINPQHIEVQILGDHTGEVHLYERDCSL 241
Db 186 DAPENLEKLAESAASREAAAFDGRVYERAVINPQHIEVQILGDSAGNIHLYERDCSL 245
QY 242 QRRHQKVEIAPAQHLDPEDRLRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNP 301
Db 246 QRRHQKVEIAPAQHLAPGLREKICADAVAFARHIGYQGAGTVEFLVDEEGNHVFIEMNP 305
QY 302 RIQVEHTVTEEVTEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGF 361
Db 306 RIQVEHTVTEEVTVQDLVKSQIMIASGATLEDLGLRQEDIHTEGAALQCRITTEDPNNGF 365
QY 362 RPDGTGTTAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRALA 421
Db 366 RPDGTGTTAYRSPGGAGVRLDGAAMLGGEITPNFDSMLVKMTCRGADFATAVARAQRALA 425
QY 422 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 481
Db 426 EFWVSGVATNIGFLRALLREEDFQHKRIATGFIADHPWLLQAPPADDEPGRILNYLDV 485
QY 482 VNKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVDTT 541
Db 486 VNKPHGLRPA-VINPVEKLPAEIKGELPRGSRDLRLQLGPEGFARALRKQDALAVDTT 544
QY 542 RDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDEL 601
Db 545 RDAHQSLLATRVRSNTLIDAARHVAKLTPELLSVEAWGGATYDVAMRFLHEDPWERLDHL 604
QY 602 REAMPNVIQMLLRGRNTVGYTPYDPSVCRAFVKEAAASSGVDFRIFDALNDVSQMRPAI 661
Db 605 REAMPNVIQMLLRGRNTVGYTPYDPSVCRAFVDEAARSGVDIFRIFDALNDVSQMRPAI 664
QY 662 DAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEELVKSGAHILAIKDMAGLLRP 721

Db 665 DAVLETNTTIAEVAMAYSGDLTNPSEKLYTLDYLYKLAEIVKSGAHVLAIKDMAGLMKP 724
QY 722 AAVTKLVLTALRREFDLPVHVHTHTDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSPSL 781
Db 725 AAATKLVTELNRNFDLPVHVHTHTDTAGGQLATYWAAAAAGADAVDGASAPLSGTTSPSL 784
QY 782 SAIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPGPTGRVYRHEIPGGQLS 841
Db 785 SAIVAAAFANTYRTDGLSLDAVGSMEPYWEAVRKLYAPFESGTPGPTGRVYQHEIPGGQLS 844
QY 842 NLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDALHLVAGVDPADFA 901
Db 845 NLRAQATALGLADRFELIEDYAAVNEMLGRPTKVTTPSSKVVGDALYLVAGVNPADFA 904
QY 902 ADPOKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHLDAD 961
Db 905 ADPOKYDIPDSVIAFLRGELGTPPGGWPEELRAKALAGRKESKDTLAPLPAEDEALL-A 962
QY 962 DSKERRNSLNRLFPKPTTEEFLEHRRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPL 1021
Db 963 DRATVRPTLDRLLFPKPAAEFAEHRRRQFGDTTKLGDAEFFYGLKEGKETVIRTADSSVPM 1022
QY 1022 LVRLDAISEPDDKGNRVNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAEFAG 1081
Db 1023 LVRLDAVGEPDEKGNRVNVVNVNGQIRPILVRDRSVESVTASVEKADPSNAGHVAAPFAG 1082
QY 1082 VVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVTVVPAATKVEGGDLIVVV 1139
Db 1083 VVTVTIEGATVKGDPVAVIEMKMEATISATDGTVDRIVLTOATKVEGGDLLLVI 1140

RESULT 7

US-10-282-122A-61950
; Sequence 61950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61950
; LENGTH: 1127
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-61950

Query Match 65.5%; Score 3791.5; DB 12; Length 1127;
Best Local Similarity 66.4%; Pred. No. 2.5e-290;
Matches 750; Conservative 140; Mismatches 229; Indels 11; Gaps 7;

QY 14 KILVANRGEIAVRAFAALETGAATVAIYPPREDRGSFHRSPASEAVRIGTEGSPVKAYLD 73
Db 4 KVLVANRGEIAIRAFAAAYELEMATVAVYPYEDRNSVHRLKADESQYIGEEGHPVRAVLS 63
QY 74 IDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRVTA 133
Db 64 VDEIVGTALACGADAIYPGYGFLSENPDLAACAAAGITFVGPSAEVLELTGDKSRAIAA 123
QY 134 AKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGGRMRFVASPDDELRLKATE 193
Db 124 ARAAGLPVLAASSPPSTSVQELLSSAAETWTFPLFVKAVAGGGGGRMRRVTDPGALAEIAEA 183
QY 194 ASREAEAAFGDGAIVYVERAVINPOHIEVOILGDHTGEVHVHLYERDCSLQRRHQKVEIAP 253
Db 184 ASREAEAFGDASVLEQAVINPRHIEVQILADTHGNVMHLYERDCSVQRRHQKVEIAP 243
QY 254 AQHLDPELDRICADAVKFCRSIGYQGACTVEFFLVDEKGNHVFIEHNPRIQVEHTVTEEV 313
Db 244 APNLDPALRERICADAVAFARSIGYTCAGTVEFFLDERGNHVFIEHNPRIQVEHTVTEEI 303
QY 314 TEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRS 373
Db 304 TDVDLVSAQLRIASGQTLLEEIGLSQDSVISRGAALQCRITTEDPANGFRPDTGRITAYRT 363
QY 374 PGAGVRLDGAAQLGGEITAHFDSMLVKMTCGSDDFETAVARAQALAEFTVSGVATNIG 433
Db 364 PGAGIRLDGGTTLGAEISAHFDSMLIKLTCGRDFTAVRRARRAVAEFRIGVSTNIP 423
QY 434 FLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEOQRILDYLDVTVNKPVGVRPKDV 493
Db 424 FLQAVLDDPDPFQAGRITTSFIEQRPOLLTARSSADRGTKILNYLDVTVNKPVGHERPSAV 483
QY 494 AAPIDKLPNIKDLPL--PRGSRDRLLKQLGPAAFARDLREQDALAVTDTTFRDAHQSLLAT 551
Db 484 -YPHDKLDPDI-DLAASPPAGSKQRLTELGPERRFAAWLRESGGVGTDTTFRDAHQSLLAT 541
QY 552 RVRSFALKPAEAEAVAKLTPELLSVEAWGGATVDVAMRFLFEDPDWDRDLDELREAMPNVIQ 611
Db 542 RVRTSGLLKVAPYIARTMPEQLLSVECWGGATVDVALRFLKEDPWERLTALREAMPNICLO 601
QY 612 MLLRGRNTVGYTYPDSVCRAVFKEAASSGVDFIFRIPDALNDVVSQMRPAIDAVLENTAV 671
Db 602 MLLRGRNTVGYTYPETVTTAFVAEATETGIDIFRIPDALNNVDSMRPAIDAVRETGSAI 661
QY 672 AEVAMAYSGDLSDPNEKLYTLDYLYKMAEEIVKSGAHILA IKDMAGLLRPAAVTKLVLTAL 731
Db 662 AEVAMSYTGDLSDPAEKLYTLDYLYKLADQIVEAGAHVLA IKDMAGLLRAPAAATLVLSAL 721
QY 732 RREFDLPVHVHTHTDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSPSLSAIVAAFAHT 791
Db 722 KSRFDLPVHVHTHTDTPGGQLATYVAAWQAGADAVDGAAPLAGTTSQPALSSIVAAAANT 781
QY 792 RRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPGPTGRVYRHEIPGGQLSNLRAQATALG 851
Db 782 EYDTGLSLPAVCDLEPYWEALRKVYAPFESGLPAPTGRVYHHEIPGGQLSNLROQAIALG 841
QY 852 LADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDALHLVAGVDPADFAADPOKYDIPD 911
Db 842 LGDRFEDIENAYAGADAILHLVKVTPSSKVVGDALALVAGVGAQDFAEDPSRYDIPD 901
QY 912 SVIAFLRGEIENPPGGWPEPLRTRALEGRSEK--APLTEVPEEQAHLDADDSKERRNS 969

Db 902 SVIGFLRGELGDPGGWPEPLRTKALQGRCPAKPEOPLT---AEDEAALAAPGAR-RQAA 957
Qy 970 LNRLFPKPTTEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAIS 1029
Db 958 LNRLFPKPTKELEHREOYQDTSGLSANQFFYGLRQGDHHRVEL-ERGVELLIGLEAIS 1016
Qy 1030 EPDDKGMNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVTVTVAE 1089
Db 1017 DPDERGMRTVMCILNGQLRPVVVRDRSIAIDVPAAEKADRANPDHIAAPFAGVTVVAEV 1076
Qy 1090 GDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVV 1139
Db 1077 GRQVEAGQTIATIEAMKMEAAVTSKSGKVARIASVSRTAQVEGGDLIMVI 1126

RESULT 8

US-10-282-122A-62684
; Sequence 62684, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62684
; LENGTH: 1127
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62684

Query Match 64.1%; Score 3712.5; DB 12; Length 1127;
Best Local Similarity 64.4%; Pred. No. 4.5e-284;
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;
Qy 12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASBAVRIGTEGSPVKAY 71
Db 2 FSKVLVANRGEIAVRAFAAYELGVGTVAVYPYEDRNSQHLKADESYQIGDIGHPVHAY 61

Qy 72 LDIDEIIGAAKVKADAIYPGYGFLSENAQLARECAENGITFIQTPTPEVLDLTGDKSRV 131
Db 62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVSPSAEVL ELAGNKSRAI 121
Qy 132 TAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFFVASPDELRLKLA 191
Db 122 AAAREAGLPVLMSSAPSASVDLSSVAAGMPFPLFVKAVAGGGGRGMRVGDIAALPEAI 181
Qy 192 TEASREAEAFDGAIVVERAVINPQHIEVQILGDHTGEVHLHYERDCSLORRHOKVVEI 251
Db 182 EAASREAESAFDPTVYLEQAVINPHIEVQILADNLGDVHLHYERDCSVQRRHOKVIEL 241
Qy 252 APAQHLDPDELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPRIOVEHTVTE 311
Db 242 APAPHLDAELRYKMCVDVAFARHIGYSCAGTVEFLDERGEYVFIEMNPRVQVEHTVTE 301
Qy 312 EVTEVDLVKAQMRLLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
Db 302 EITDVDLVASQLRIAAGETLEQLGLRQEDIAHPGAALQCRITTEDPANGFRPDTGRISAL 361
Qy 372 RSPGGAGVRLDGAALGGEITAHFDSMLVKVTCRGSDFETAVARAQALAEFTVSGVATN 431
Db 362 RTAGGAGVRLDGSNLCAGAEISPYFDSMLVKLTCRGRDLPTAVSRARRAIAEFRIRGVSTN 421
Qy 432 IGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDVLADVTNKPCHGVPRPK 491
Db 422 IPFLQAVLDDPFRAGRVTTSFIDERPQLLTARASADRGTGKILNLFADVTNPNPYGSRPS 481
Qy 492 DVAAPIDKLPNTKDLP-----PRGSRDLKQLGPAAAFARDLREQDALAVTDTTFRDAHQS 547
Db 482 TI-YPDDKLP---DLDLRAAPPAGSKORLVKLGPEGFARWLRESAAVGVTDTTTFRDAHQS 537
Qy 548 LLATRVRSFALKPAAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDEPDWRDLDELREAMPN 607
Db 538 LLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLRAAMPN 597
Qy 608 VNIQMLLRGNTVGYTPYSDSVCRAFKVAEASSGVDFIRFDALNDVSMRPAIDAVLET 667
Db 598 ICLQMLLRGNTVGYTPYPEIVTSAFVQEATATGIDIFRIFDALNNIESMRPAIDAVRET 657
Qy 668 NTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIDMAGLLRPAAVTKL 727
Db 658 GSAIAEVAMCYTGDLPDPEQLYTLDYLYLKLAEQIVDAGAHVLAIDMAGLLRPPAAQRL 717
Qy 728 VTALRREFDLPVHVHTHTAGGOLATYFAAAAQAGADAVDGASAPLSGTTSPSLSAIVAA 787
Db 718 VSALRSRFDLPVHLHTHTDTPGGQLASYVAAWHAGADAVDGAAPLAGTTSPALSSIVAA 777
Qy 788 FAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQA 847
Db 778 AAHTEYDTGLSLSAVCALEPYWEALRVYAPFESGLPGPTGRVYRHEIPGGQLSNLRQQA 837
Qy 848 TALGLADRFELIEDNYAAVNEMLRPTKVTTPSSKWVGDALHLVAGVDPADFAADPOKY 907
Db 838 IALGLGDRFEEIEEAYAGADRVGLRVKVTPTSKVWGDALALVAGVSADEFASDPARF 897
Qy 908 DIPDSVIAFLRGEIGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHLDADDSKERR 967
Db 898 GIPESVLGFLRGELGDPGGWPEPLRTAALAGGAAR-PTAQLAADDEIALSSVGAK-RQ 955
Qy 968 NSLNRLLPKPTTEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
Db 956 ATLNRLLPSPPTKEFNEHREAYGDTSQLSANQFFYGLRQGEHHRVKL-ERGVELLIGLEA 1014
Qy 1028 ISEPDDKGMNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGVTVTV 1087
Db 1015 ISEPDERGMRTVMCILNGQLRPVLVRDRSIAVPAAEKADRGNPNHIAAPFAGVTVTV 1074
Qy 1088 AEGDEVKAGDAVAIIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db 1075 CVGERVAGQTIATIEAMKMEAPITAPVAGTVERVAVSDTAQVEGGDLIVVVS 1127


```

; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52777
; LENGTH: 1144
; TYPE: PR
; ORGANISM: Clostridium botulinum
US-10-282-122A-52777

Query Match      44.1%; Score 2552; DB 12; Length 1144;
Best Local Similarity 45.7%; Pred. No. 2.8e-192;
Matches 526; Conservative 215; Mismatches 372; Indels 38; Gaps 14;

QY      12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEFASAEAVRIGTEGSPVKAY 71
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      4 FKRVLVANRGEIAIRVFRACHELGIRTVAIYSEEDKFSLFRTKADAEAYLIGNKNGPIDAY 63

QY      72 LDIDEIIGAAKVKKADAIYPGYGFLSENAQLARECAENGITFCPTPEVLDTGDKSRVAV 131
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      64 LNIEEIIQLALKKGVDIAHPGYGFLAENSEFARKCREAGIEFICGPTAEMMEKLGDKIKSK 123

QY      132 TAACKAGLPVL-ABSTPSKNIDEIVKSAEQTYPIFVKAVAGGGGRGMRFVASPDDELRLK 190
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      124 IVAEKAGVPTIPGVQKPIKSEKEALEFARYCGYPIMLKAAAGGGGRGMRIVRTEELISS 183

QY      191 ATEASREAEAFGDCAVYVERAVINPQHIEVQILGDHTGTEVVHLVERDCSLQRRHQKVVE 250
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      184 FKSANEAKKAFGIDDDIFIEKYLENPKHIEVQILGDKHGNIVHLVERDCSIQRRHQKVIE 243

QY      251 IAPAOHLDPEDRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEPNRIQVEHTVT 310
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      244 FTPAFALPKKKRREIEICNDALKIAKTGVYRSAGTLEFLVDTTGNHVFIEPNRIQVEHTVT 303

QY      311 EEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDGTGT 367
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      304 EMITGIDIVQSQILIAEGYKLDSEEVGIKSQESIQTRGAIQCRVTTEDPSNNFAPDTGK 363

QY      368 ITAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVQWTCRGSDFEFAVARAQRALAEFTVS 426
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      364 IEDYRTSGFGIRLDGNGGFTGVSIPSYDLSLVKTTWSRTTFNDAIRKSIRAIKEFKID 423

QY      427 GVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYLDADVTVNKPH 486
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      424 GVKTNIGFLINVLNHEQFRKGQCDTNFIEKNPELFDITSKTDDEVRILKFIGEKVNVETH 483

QY      487 GVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTTFRDAHQ 546
```

```

Db      484 GIK-KDFDVP--TIPIVDEGLSLKGTQKILDEKPGELVSWIKTQNKLLLTDTTNRDAHQ 540
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      547 SLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWDLRLDELREAMP 606
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      541 SLMATRMSVDMFKIAKAQSVLQKDLFSMEMWGGATFDVAYRFLKESPWTRLEELRSIP 600
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      607 NVNIQMLLRNTVGYTPYPSVVCRAFVKEAASSGVDFRIFDALNDVSQMRPAIDAVLE 666
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      601 NVLFQMLIRGANAVGYKNYPDNVIRKFIKQSAADSGIDVFRIFDLSNLWLGMEVATDEVLK 660
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      667 TMTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIKDMAGLLRPAAVTK 726
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      661 QN-KVAETCMCYTGDILEEYRDYSLQYYVDLAKDIEKTGAHILGIKDMSALLKPYAAVK 719
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      727 LVTALRRRFDLPVHVHTHTDTAGGQATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVA 786
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      720 LKALKNEISIPILHHTHTDTTNGVATVLMAAHAGVDIVDTAFNSMSGTSLTSPALNSIVA 779
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      787 AFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGQLSNLRAQ 846
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      780 ALENTDRETGLDLDTMQKLSDYWSAVRPVYSQFESGLKSGSAEIKYEIPGGQYSLNLFQ 839
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      847 ATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDIALHLVGAGVDPADFAADPOK 906
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      840 VESFGLGHKFEVEMKMYKKNEMLGDIIKVTPSSKVVGDIAIFMVKNDLTPENIYEKAEK 899
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      907 YDIPDSVIAFLRGELGNPPGGWPELRTALEGRSEKAPLTVPEE----- 953
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      900 MAFPDSAVSYFKGMGQPMGGFPEKLQKLVKGED----PITCRPGEMLPPEDFEKIREH 955
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      954 --EQAHLDADDKSKERNNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETL 1011
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      956 LKEKHDLDATEN---DIISYALYFEVFDKYLDLFLKEYGDLSHMGSDVFFHGLYEGETAE 1011
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1012 IRLPDVRTPLLVRLDAISEPDDKGMNVVNVANVNGQIRPMRVDRS---VESVTATAEK-A 1067
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1012 IELQEGKT-FIVQLSEIGKVDSEGNRAVVFEINGNRREIRIKDKSSLMAQNITSNSTKMA 1070
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1068 DSSNKGHVAAPPAG-VVTVTVAEAGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVP 1126
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1071 DPANKKHIGSSIPGTVIKVLVKNKGEIKEGDSLIVIAMKMETNIVASLSGVVGSLLVKE 1130
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1127 ATKVEGGDLIV 1137
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1131 GDQVRSGQLLL 1141
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```

RESULT 11
US-10-282-122A-46866
; Sequence 46866, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```



```

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46866
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-46866

Query Match      43.6%; Score 2526; DB 12; Length 1148;
Best Local Similarity 47.2%; Pred. No. 3.2e-190;
Matches 542; Conservative 178; Mismatches 402; Indels 26; Gaps 13;

QY 9 LPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV 68
Db 4 LQRIQKVLVANRGEIAVRFACSELGLKTVAIYSKEDSGSYHRYKADESYLVGEGKKPI 63

QY 69 KAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS 128
Db 64 DAYLDIEGIIIEIAKSNHVDIAHPGYGFLSENIQPAKCEEEGIIFIGPKSKHLDPMFGDKV 123

QY 129 RAVTAAKKAGLPVLAEST-PSKNIDEIVKSAEGOTYPIFVKAVAGGGGRGMRFVAPDEL 187
Db 124 KARTQAQLAIPVPGSDGPDVDSLEEVEKFAEKYDYPYIIKASLGGGGRGMRIVRTSEEL 183

QY 188 RKLATEASREAEAFGDGAVYVERAVINPQHIEVQILGDHTGEVHVLYERDCSLQRRHQK 247
Db 184 RESYNRAKSEAKAAGFNDEYVEKFEKPKHIEVQILADEEGNVVHLYERDCSVQRRHQK 243

QY 248 VVEIAPAQHLDPEDRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEPNRIQVEH 307
Db 244 VVEIAPSVLSDDLQRICEAAVKLTKNVNYLNAGTVEFLVKD-DNFYFIEVNPVRVQVEH 302

QY 308 TVTEEVTEVDLVKAQMRILAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNGRFPD 364
Db 303 TITEMITGVDIVQSQIILADGHALSHKMGVGPKEBEEVVHGFQAIQSRVTTEDPLNNFMPD 362

QY 365 TGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEF 423
Db 363 TGMIMAYRSGGGFVRLDTGNSFQGAVIAPYDSSLVKVTTWALTFEQAAKWERNLKEF 422

QY 424 TVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDADVTVN 483
Db 423 RIRGIKTNIPIFLENVVKHKNFLSGEYDTSFIDASPELFLFKPKRDKRGTKMLNYIGTVTN 482

QY 484 KPHGVRPKDVAA-PIDKLPNI-KDLPFRGSRDRUKQLGPAAAFARDLREQDALAVTDTTF 541
Db 483 GFPGVGKKEKPIFPDARIPNVLHSEPIQNGTKQLDBERGADGLVKVQDQKRVLTTDTTF 542

QY 542 RDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDL 601
Db 543 RDAHQSLLATRIKTLHQIAEPTARMLPNLFSAEWGGATFDVAYRFLKEDPWERLLDL 602

QY 602 REAMPNVNIQMLLRGRNTVGYTPYSDVCRAFKVEAASSGVDFIRIFDALNDVDSQMRPAI 661
Db 603 REKMPNVLFQMLLRSSNAVGYKNYPDNLIOKFVECSAQAGIDVFRIFDLSLNWVEGMRVAI 662
```

```

QY 662 DAVLETTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEBEIVKSGAHILAIKDMAGLLRP 721
Db 663 DAVRDIG-KIAEATMCTGDIHDPMRSKYDLNYYKNLAKELEVSGAHILGIKDMAGLLKP 721

QY 722 AAVTKLVTLRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSL 781
Db 722 NAAAYDLVSALKETVSIPIHLHTHTDSNGILTYTKAIEAGVDIVDVAVSSMAGQTSQPSA 781

QY 782 SAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLS 841
Db 782 NTLYYALGGNERQPDVNIDSLEKLSHYWEDVRKYAPFESGNNAPHTEVYMHMPGGQYS 841

QY 842 NLRAQATALGLADRFELIEDNYAAVNEMGLRPTKVTWTPSSKVVGDLALHLVAGVDPADFA 901
Db 842 NLQQQAKVVVGLGDRFDEVKVMYRRVNDMFGDIVKVTWTPSSKVVGDMALFMVQNHLEQDVL 901

QY 902 ADPQKYDIPDSVIAFLRGEIENPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHLDAD 961
Db 902 ERGHSMDPFGSVVEMFSGDLGQPYGPFKKLQEIIL----KGKEPLTVRPGELLEVPVDFD 957

QY 962 DSKERR-NSLNR-----LLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRET 1010
Db 958 ALKEELFHKLGREVTMFDVVAYALYPKVFMDYKVAELYGNVSVLDTPTFFYGMRLGEEI 1017

QY 1011 LIRLPDVRTPLVRLDAISEPDDKGMNVVNVANVNGQIRPMRVDRSVESVTATAEKADSS 1070
Db 1018 DVEIEQGKT-LMVKLVSIGELQPDGNRVLYLEFNGQPREIIVKDESVKATVAQRVKGNRE 1076

QY 1071 NKGHVAAPFAG-VVTVTVAEDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVAATK 1129
Db 1077 NPNHISATMPGTIVKVVVKEGDEVKKGDSMAITEAMKMETTVQAPFNGKVKVYVNDGDA 1136

QY 1130 VEGGDLIV 1137
Db 1137 IQTGDLII 1144
```

```

RESULT 12
US-10-369-493-17338
; Sequence 17338, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17338
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-10-369-493-17338

Query Match      43.6%; Score 2525.5; DB 15; Length 1150;
Best Local Similarity 46.8%; Pred. No. 3.5e-190;
Matches 539; Conservative 185; Mismatches 400; Indels 27; Gaps 13;
```

```

QY 9 LPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPV 68
Db 4 LKNIKKVLVANRGEIAIRIFRACTELHIRTVALYISKEDTGAYHRYKADEAYLVGEGKKPI 63

QY 69 KAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS 128
Db 64 EAYLDIEGIIIEIAKRHGVDIAHPGYGFLSENIEBPAKRCHEEGIIFIGPELEHLVFMFGDKV 123
```

QY 129 RAVTAAKKAGLPVLAEST-PSKNIDEIVKSAEGOTYPIFVKAVAGGGGRMRFVASPDDEL 187
Db 124 QAREQAIKANLPVPGSDGVPSSLEDVKAFADKHGYPFIKAALGGGRMVRVRSNDV 183
QY 188 RKLATEASREAEAAFGDGAUVVERAVINPQHIEVQILGDHTGEVHVLYERDCSLQRRHQK 247
Db 184 QESYERAKSEAKAAGFNDEVVEKFIENPKHIEVQILADKHGNTLHLVERDCSVQRRHQK 243
QY 248 VVEIAPAQHLDPELRDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFLEMPRIQVEH 307
Db 244 VVEAPSVLSLSEDRERICQAAVQLAENNVYVAGTVEFLVDREGNFYFIEVNPRIQVEH 303
QY 308 TVTEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNGFERPD 364
Db 304 TITEMVTGIDIVQSOLFIADEGHLHGDRIGIPKQEEIVCHGYAIOQSRVTTEDEPSNGFLPD 363
QY 365 TGITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEF 423
Db 364 TGRINAYRSGGGFGVRLDAGNGFQGAVIDTPPYDSLLVKVSTWALTTFEGAAAKMLNREF 423
QY 424 TVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDVTVN 483
Db 424 RIRGIKTNIAFLENVVQHRQFSLGEYNTSFIDQTPFLFVPPKRDGRTKMLSFIGETIVN 483
QY 484 KPHGVPRKDVAAPIDK--LPNIK-DLPLPRGSRDLKQLGPAAFARDLREQDALAVDTT 540
Db 484 GYPGLE-KTKKPVFDKPPVPKULKLEPIPDGTQKQILDQHGPEGLAKVWKEKHVLLTDTT 542
QY 541 FRDAHQSLLATRVRSFALKPAABAAVAKLTPELLSVEANGGATYDVAMRFLFEDPDWRDLDE 600
Db 543 FRDAHQSLLATRVRTHDLKQIAEPTARLLPNLFSMEMWGGATFDVAMRFLHEDDPWERLLI 602
QY 601 LREAMPNVNIQMLLRGRNTVGTPYPDSVCRAVFVKEAASSGVDIFRIFDALNDVSMRPA 660
Db 603 LRKKAPNVLFQMLLRASNNAVGYKNYPDNLIREFVDKSNAGIDVFRIFDSLNVWEGMKLA 662
QY 661 IDAVJETNTAAEVAEMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDMAGLLR 720
Db 663 IEAVGEAN-KIAEATICYTGDIILSSRPKPYDLAYYKLAKELEAAGAHILGIDMAGLLK 721
QY 721 PAAVTKLVALTALRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPISGTTSQPS 780
Db 722 PEAAVQLVLAELKDTVTIPVHLHTHTSGNGIFTYARAIEAGVDIVDVAVSSMAGLTSQPS 781
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVYRHEIPGGQL 840
Db 782 ANSLYALADSERQPNVNITALEQAEFWEEETRFKYAGFESGMNAPHTEVYEHMPGGQY 841
QY 841 SNLRAQATALGLADRFELEDNYAAVNEMLGRPTKVTSPSSKVVGDLALHLVGAGVDPADF 900
Db 842 SNLQOQAKAVGLGHRWNEVKMYRTVNDMFGDVVKVTPSSKVVGDMALYVQNDLTEEV 901
QY 901 AADPQKYDIPDPSVIAFLRGELNPPGGWPEPLRTRALEGSRSEGKAPLTEVPPEEQAHLDA 960
Db 902 YENGHKLDPDPSWVEFFEGQLGQPYQGPPKKLQEIILKGRK---PITNRPGENMEPIQF 957
QY 961 DDSKER-RNSLNR-----LLPPKPTEEFLEHRRRRFGNTSALDDREFFYGLVEGRE 1009
Db 958 EAIKEELYNKLDROQVTSHDILSYALYPKVFMEFERFRQTFGDSVSLDTPTTFYGLRPGEE 1017
QY 1010 TLIRLPDVRTPLLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADS 1069
Db 1018 IEVEIEQGT-LIVKFISLSKPQDDGNRIYVYFELNGQPREVLIKDSVKTSIISRPKADK 1076
QY 1070 SNKGHVAAFPAG-VVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAAT 1128
Db 1077 SNPNHIGASMPGTVVKALVEKGDVKVQGDHLMITEAMKMETIVQAPFDGEVVALHKDGD 1136
QY 1129 KVEGGDLIVVV 1139
Db 1137 AIQTGDLILIEV 1147

RESULT 13
US-10-369-493-23112
; Sequence 23112, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23112
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23112

Query Match 43.6%; Score 2524; DB 15; Length 1148;
Best Local Similarity 46.0%; Pred. No. 4.6e-190;
Matches 533; Conservative 186; Mismatches 392; Indels 48; Gaps 13;

QY 11 AFKKTILVANRGEIAVRAFAALETGAATVAIYPREDRGSHFRSFASEAVRIGTEGSPVKA 70
Db 5 SIQKVLVANRGEIAIRIFRACTELNIRTVAVYSKEDSGSYHRYKADEAYLVGEGKKPIDA 64
QY 71 YLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA 130
Db 65 YLDIEGLIDIAKRNKVDAIHPGYGFLSENIHFARRCCEEGIVFIGPKSEHLDMFGDKVKA 124
QY 131 VTAACKAGLPVLAEST-PSKNIDEIVKSAEGOTYPIFVKAVAGGGGRMRFVASPDDELRK 189
Db 125 REQAEKAGIPVIPGSDGPAETLEAVEQFGQANGYPIIITKASLGGGGRMIRVSESEVKE 184
QY 190 LATEASREAEAAFGDGAUVVERAVINPQHIEVQILGDHTGEVHVLYERDCSLQRRHQKV 249
Db 185 AYERAKSEAKAAGFNDEVVEKFIENPKHIEVQVIGDKQGNVHLFERDCSVQRRHQKVI 244
QY 250 EIAPAQHLDPELRDRICADAVKFCRSIGYOGAGTVEFLVDEKGNHVFLEMPRIQVEHTV 309
Db 245 EVAPSVLSLPELRDQICEAAVALAKNVYINAGTVEFLV-ANNEFYFIEVNPRVQVEHTI 303
QY 310 TEEVTEVDLVKAQMRLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFRPDTG 366
Db 304 TEMITGVDIVQTQILVAQGHSLHKKVNIPEQKIDFTTGYAIOQSRVTTEDEPQNDMPDTG 363
QY 367 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTV 425
Db 364 KIMAYRSGGGFGVRLDTGNSFOGAVITPYDSLLVKSLLVLTWALTTFEQAAAAMVRNLQEFRI 423
QY 426 SGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDVTVN-- 483
Db 424 RGIKTNIFFLENVAKHEKFLTGQYDTSFIDTTPELFNFPPKQKDRGTKMLTYIGNVTVNGF 483
QY 484 -----KPHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAV 536
Db 484 PGIGKKEKPAFDKPLGVKXVDVQQP-----ARGTKQILDKEKGAEGLANVWKEQKSVLL 536
QY 537 TDTTFRDAHQSLLATRVRSFALKPAABAAVAKLTPELLSVEANGGATYDVAMRFLFEDPDW 596
Db 537 TDTTFRDAHQSLLATRIRSHDLKKIANPTAALWPELFSMEMWGGATFDVAYRFLKEDPWK 596
QY 597 RLDELREAMPNVNIQMLLRGRNTVGTPYPDSVCRAVFVKEAASSGVDIFRIFDALNDVSO 656
Db 597 RLEDLRKEVPNTLQMLLRSSNAVGYTNPYDNVKEFVKQSAQSGIDVFRIFDSLNVWVK 656
QY 657 MRPAIDAVLETNATAEVAEMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDMA 716

Db 657 MTLAIDAVRTG-KVAEAAICYTGDIIDKNRTKYDLAYTSMKELEAAGAHILGIKDMA 715

Qy 717 GLLRPAAVTKLVTALRREFDLPVHVHTHTAGGLATYFAAAQAGADAVDGSAPLSGTT 776

Db 716 GLLKPOAAEYLVSAKLETIDIPVHLHTDTSNGIYMYAKAVEAGVDIIDVAVSSMAGLT 775

Qy 777 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 836

Db 776 SQPSASGFYHAMEGNDRRPEMNVQVVELLSQYWESVRKYYSFESGKSPHTEIYHEMP 835

Qy 837 GGQLSNLRAQATALGADRFELEDNYAAVNMELGRPTKVTTPSSKWVGDLALHLVGAGVD 896

Db 836 GGQYSNLQQQAKGVGLGDRWNEVKEMYRRVNDMFGDIVKVTTPSSKWVGDMALYMVQNNLT 895

Qy 897 PADFAADPQKYDIPDSVIAFLRGELGNPPGWPPEPLRTRALRSEKSEKAPLTVPEE--- 953

Db 896 EKDVYKESLDFPDSVVELFKGNIGQPHGPFPEKLQKLILKQOE---PITVRPGELLE 951

Qy 954 -----EQAHLDDADDSKERRNSLNRLLPKPTPEEFLEHRRRFGNTSALDDREFF 1001

Db 952 PVSFEAIKQEFKEQHNLISD---QDAVAYALYPKVFTDYVKTESYGDISVLDTPTEFF 1007

Qy 1002 YGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMNNVANVNGQIRPMVRDRSYESVT 1061

Db 1008 YGMTLGEIEIEVEIERGKT-LIVKLISIGBPQPDATRVVYFELNGQPREVWIKDESIXSV 1066

Qy 1062 ATAEKADSSNKGHVAAPFAGVTVTVAE-CDEVKAGDAVAIIIEAMKMEATITASVDGKID 1120

Db 1067 QERLKADRTNPSTHAASMPGTVIKVLAEAGTKVNGKGDHLMINEAMKMETTVQAPPSGTIK 1126

Qy 1121 RVVVPAAATKVEGGDLIVVV 1139

Db 1127 QVHVKNGEPIQTGDLLEI 1145

RESULT 14

US-10-282-122A-60488

; Sequence 60488, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60488

; LENGTH: 1146

; TYPE: PRT

; ORGANISM: Listeria monocytogenes

US-10-282-122A-60488

Query Match 43.1%; Score 2493; DB 12; Length 1146;

Best Local Similarity 45.8%; Pred No. 1.3e-187;

Matches 526; Conservative 185; Mismatches 408; Indels 30; Gaps 11;

Qy 13 KKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVRIGTEGSPVKAYL 72

Db 5 KKVLVANRGEIAIRVMRACTELIKIKTVAIYSQEDTGSFPHRYKSDAYLVGAGKKPIDAYL 64

Qy 73 DIDEIIGAARKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVT 132

Db 65 DIENIIIEAKESGADAIHPGYGFLSENIIEFARRCEQEGIIIFVGPKSXHLDMFGDKIKAKE 124

Qy 133 AAKKAGLPVLAEST-PSKNIIDEIVKSAEQTYPIFVKAVAGGGGRGMRFVASPDELKLA 191

Db 125 QALLADIPIVPGSNPVGAGIKEVEEFGKNGYPLMIKASLGGGGRGMRVVESEKHVESF 184

Qy 192 TEASREAAEFGDGAVYVERAVINPOHIEVOILGDHTGEVVHLYERDDCSLQRRHQKVVEI 251

Db 185 ERASSEAKAAGNDEVYVEKVMNPKHIEVQILGDTHGNIHVFERDCSIQRRHQKVVEV 244

Qy 252 APAQHLDPELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVEIEMNPRIQVEHTVTE 311

Db 245 APCNAITSELNRICDAAVKLMKNVDYINAGTVEFLV-EGDDFYFIEVNPVQVEHTITE 303

Qy 312 EVTEVDLVKAQVRLAAGATLKEGLT---QDKIKTHGAALQCRITTEDPNNGFRPDTGTI 368

Db 304 MITGIDIVQSOLFADGYALHDQLVAIPKQEDIIHGSAIQSRITTEDPLNNFMPDTGRV 363

Qy 369 TAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSG 427

Db 364 DTYRSTGGFGVRLDAGNGFGQTVVTFYDLSLVKLCTWGMTFEQATRKMRNLIEFRIRG 423

Qy 428 VATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADVTVNKPHG 487

Db 424 VKTNIPFLNVVRHDPFASGNNTSFDITTPELFKFPHIRDRTKTLRYIGNVTVNGFPFG 483

Qy 488 VRPKDVAAPIDKLPNIKDL-PRGRSRDLKQLGPAAFARDLREEQDALAVTDTTFRD 543

Db 484 IKHRD--KPVYAEPRIPKIPYGSQISPGTKQILDAKGPEGVVDVWVKQEBVLLTDTTLRD 541

Qy 544 AHQSLLATRVRSFALKPAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWDRDELRE 603

Db 542 AHQSLLATRVRSKDIQVADAMAHLLPNMFSEFMWGGATFDVAYRFLNEDPWVRLETLRK 601

Qy 604 AMPNVNIQMLLRGRNTVGYTPYDPSVCRAFKVEAASSGVDFIRIFDALNDVVSQMRPAIDA 663

Db 602 QIPNVMFQMLLRGANAVGYKNYPDNVIREFVKQSAQSGVDVFRVFDLSLWIKGMEVSIDA 661

Qy 664 VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYVLMKMAEIEIVKSGAHILAIDMAGLLRPAA 723

Db 662 VREAG-KVVEAICYTGIDIDDDTRTKYITIDYVKMAKELVAQGTGTHILGIKDMAGLLKPQA 720

Qy 724 VTKLVTLRREFDLPVHVHTHTTAGGQLATYFAAAQAGADAVDGSAPLSGTSQPSLSA 783

Db 721 AYRLIGELKDTVDVPIHLHTDTSNGIYTYAAAVSAGVDIVDVASSAMSGATSPSMTG 780

Qy 784 IVAAFATRRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQLSNL 843

Db 781 LYYGLVNGNRQTNLDAQNSQLINHYWEDVRHYKDFDNLNLSPTQTEVIHEMPGGQYTNL 840

Search completed: March 24, 2004, 22:50:25
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:36:06 ; Search time 64 Seconds
(without alignments)
5032.877 Million cell updates/sec

Title: US-10-045-072-2
Perfect score: 5788
Sequence: 1 MSTHTSSTLPFAKKILVANR.....RVVVPAAATKVEGDLIVVVS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5788	100.0	1140	4	AAB67129 Coryneb
2	5788	100.0	1140	4	AAG90511 C glutami
3	5788	100.0	1140	5	AAE25601 Coryneb
4	5788	100.0	1140	5	AAU98053 Coryneb
5	5788	100.0	1140	7	ABU10426 Coryneb
6	5784	99.9	1141	3	AAB01436 Pyruvate
7	5782	99.9	1140	2	AAW93971 C. glutam
8	5780	99.9	1140	4	AAG93249 C glutami
9	5759	99.5	1140	5	AAU98052 Coryneb
10	5759	99.5	1157	5	AAU98050 Coryneb
11	5302.5	91.6	1139	4	AAB83180 Coryneb
12	4647.5	80.3	1141	6	ABU25961 Protein e
13	3791.5	65.5	1127	6	ABU34026 Protein e
14	3712.5	64.1	1127	6	ABU36886 Protein e
15	3712.5	64.1	1127	6	ABU34760 Protein e
16	3483	60.2	1124	7	ADB74261 Mycobacte
17	2621	45.3	532	4	AAB79302 Coryneb
18	2552	44.1	1144	6	ABU24853 Protein e
19	2526	43.6	1148	6	ABU18942 Protein e
20	2525.5	43.6	1178	7	ADE62415 Rat Prote
21	2525.5	43.6	1178	7	ADE62419 Rat Prote
22	2524	43.6	1148	4	AAU00511 Bacillus
23	2520.5	43.5	1178	7	ADE62421 Human Pro
24	2520.5	43.5	1178	7	ADE62417 Human Pro
25	2493	43.1	1146	5	ABB47612 Listeria

26	2493	43.1	1146	6	ABU32564	Abu32564 Protein e
27	2491.5	43.0	1144	6	ABU24164	Abu24164 Protein e
28	2488	43.0	1147	4	AAU33972	Aau33972 Staphyloc
29	2488	43.0	1150	6	ABU16467	Abu16467 Protein e
30	2485	42.9	1156	6	ABM73249	Abm73249 Staphyloc
31	2472.5	42.7	1142	4	AAU35213	Aau35213 Enterococ
32	2464.5	42.6	1142	6	ABU30018	Abu30018 Protein e
33	2464	42.6	1151	4	AAG82677	Aag82677 S. epider
34	2464	42.6	1154	5	ABP38583	Abp38583 Staphyloc
35	2463.5	42.6	1142	6	ABU14604	Abu14604 Protein e
36	2455	42.4	1144	6	ADB08196	Adb08196 Alloicoc
37	2441	42.2	1181	4	ABB66604	Abb66604 Drosophil
38	2441	42.2	1181	4	ABB66605	Abb66605 Drosophil
39	2441	42.2	1181	4	ABB58211	Abb58211 Drosophil
40	2433.5	42.0	1196	4	ABB67309	Abb67309 Drosophil
41	2405	41.6	1136	5	ABB53980	Abb53980 Lactococc
42	2375	41.0	1143	6	ABU24993	Abu24993 Protein e
43	2357	40.7	461	4	AAE79303	Aab79303 Coryneb
44	2299	39.7	1073	4	AAU36768	Aau36768 Staphyloc
45	2174	37.6	973	7	ADC95183	Adc95183 E. faeciu

ALIGNMENTS

RESULT 1
AAB67129
ID AAB67129 standard; protein; 1140 AA.
XX

AC AAB67129;
XX
DT 12-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum pyruvate carboxylase.
XX

KW Pyruvate carboxylase; anaplerotic pathway; industrial fermentation;
KW oxaloacetate.
XX

OS Corynebacterium glutamicum.
XX

PN US6171833-B1.
XX

PD 09-JAN-2001.
XX

PF 23-DEC-1998; 98US-00220081.
XX

PR 23-DEC-1998; 98US-00220081.
XX

PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX

PI Sinskey AJ, Lessard PA, Willis LB;
XX

DR WPI; 2001-122330/13.
XX

DR N-PSDB; AAF32165.
XX

PT Novel nucleic acid encoding pyruvate carboxylase from Corynebacterium
glutamicum, for replenishing oxaloacetate consumed during lysine and
glutamic acid production in industrial fermentations.
XX

PS Claim 1; Col 31-36; 29pp; English.
XX

CC The present invention provides the protein and coding sequences of the
Corynebacterium glutamicum pyruvate carboxylase protein. This is an
enzyme in the anaplerotic pathway. It can be used in the replenishment of
oxaloacetate consumed during lysine and glutamic acid production in
industrial fermentation
XX

SQ Sequence 1140 AA;
XX

Query Match 100.0%; Score 5788; DB 4; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

app. cont.

Qy 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR 60
Db 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR 60
Qy 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120
Qy 121 LDLTGDKSRVATAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180
Db 121 LDLTGDKSRVATAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180
Qy 181 VASPDRLKRLATEASREAEAFDGDGAVYVERAVINPQHIEVQILGHTGEVHLYERDCS 240
Db 181 VASPDRLKRLATEASREAEAFDGDGAVYVERAVINPQHIEVQILGHTGEVHLYERDCS 240
Qy 241 LQRRHQKVEIAPAOHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEVN 300
Db 241 LQRRHQKVEIAPAOHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEVN 300
Qy 301 PRIQVEHTVTEEVTLVKQAMRLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTLVKQAMRLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 360
Qy 361 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
Db 361 FRPDTGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
Qy 421 ABFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
Db 421 ABFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480
Qy 481 TVNKPVGVRPKOVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKOVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
Qy 541 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWDLDE 600
Db 541 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWDLDE 600
Qy 601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFYKEAASSGVDFRIFDALNDVVSQMRPA 660
Db 601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFYKEAASSGVDFRIFDALNDVVSQMRPA 660
Qy 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDVYVKMAEEIVKSGAHLIAIKDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDVYVKMAEEIVKSGAHLIAIKDMAGLLR 720
Qy 721 PAAVTKLVTLRREFDLPVHVHTHTDAGGQATYFAAAQAGADAVDASAPLSGTTSQPS 780
Db 721 PAAVTKLVTLRREFDLPVHVHTHTDAGGQATYFAAAQAGADAVDASAPLSGTTSQPS 780
Qy 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRYRHEIPGGQL 840
Db 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRYRHEIPGGQL 840
Qy 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVDLALHLVAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVDLALHLVAGVDPADF 900
Qy 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTVPEVEEQAHLDA 960
Db 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTVPEVEEQAHLDA 960
Qy 961 DDSKERRNSLNRLLFPKPTTEFLHRRRFGNTSALDDREFFYGLVEGRETLRLPDVRTP 1020
Db 961 DDSKERRNSLNRLLFPKPTTEFLHRRRFGNTSALDDREFFYGLVEGRETLRLPDVRTP 1020
Qy 1021 LLVRLDAISEPDDKGMNRNVVANVNGQIRPMRVDRDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db 1021 LLVRLDAISEPDDKGMNRNVVANVNGQIRPMRVDRDRSVESVTATAEKADSSNKGHVAAPFA 1080
Qy 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140

Db 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
RESULT 2
AAG90511
ID AAG90511 standard; protein; 1140 AA.
XX AAG90511;
AC AAG90511;
XX 26-SEP-2001 (first entry)
DT 26-SEP-2001 (first entry)
XX C glutamicum protein fragment SEQ ID NO: 4265.
DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
KW Corynebacterium glutamicum.
XX OS EP1108790-A2.
PN 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
PF 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
PI WPI; 2001-376931/40.
DR N-PSDB; AAH65730.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX Claim 17; SEQ ID NO 4265; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 1140 AA;
Query Match 100.0%; Score 5788; DB 4; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR 60
Db 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAVR 60
Qy 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120
Qy 121 LDLTGDKSRVATAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180
Db 121 LDLTGDKSRVATAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRF 180

QY 181 VASPELRLKLAATEASREAEAAFGDGA VYVERAVINPQHIEVOILGDHTGEVVHLYERDCS 240
Db 181 VASPELRLKLAATEASREAEAAFGDGA VYVERAVINPQHIEVOILGDHTGEVVHLYERDCS 240
QY 241 LORRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEWN 300
Db 241 LORRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEWN 300
QY 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
QY 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
Db 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREODALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREODALAVTDTT 540
QY 541 ERDAHQSLLATVRVSFALKPAABAVAKLTPELLSVEAWGGATVDVAMRFLFEDPWRDLDE 600
Db 541 ERDAHQSLLATVRVSFALKPAABAVAKLTPELLSVEAWGGATVDVAMRFLFEDPWRDLDE 600
QY 601 LREAMPNVNIQMLJRGNTVGYTPYDPSVCRAFVKEAAASSGVDFRIFDALNDVSNMRPA 660
Db 601 LREAMPNVNIQMLJRGNTVGYTPYDPSVCRAFVKEAAASSGVDFRIFDALNDVSNMRPA 660
QY 661 IDAVLETNTAEVAVAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLR 720
Db 661 IDAVLETNTAEVAVAYSGDLSDPNEKLYTLDYLLKMAEIVKSGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTLALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSPS 780
Db 721 PAAVTKLVTLALRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSPS 780
QY 781 LSAIVAFAFHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db 781 LSAIVAFAFHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
QY 841 SNLRAQATALGLADRPELIEDNYAAVNEMLGRPTKVTPTSSKVVGDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRPELIEDNYAAVNEMLGRPTKVTPTSSKVVGDLALHLVGAGVDPADF 900
QY 901 AADPKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHILDA 960
Db 901 AADPKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHILDA 960
QY 961 DDSKERRNSLNLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db 961 DDSKERRNSLNLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
QY 1021 LLVRLDAISEPDDKGMNRNVVANVNGQIRPNRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db 1021 LLVRLDAISEPDDKGMNRNVVANVNGQIRPNRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1081 GWTVTVAEGDEVKAGDAVAIIAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
Db 1081 GWTVTVAEGDEVKAGDAVAIIAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140

RESULT 3
AAE25601
ID AAE25601 standard; protein; 1140. AA.
XX
AC AAE25601;
XX
DT 04-NOV-2002 (first entry)

XX
DE
XX
KW
KW
XX
OS
XX
FH
FT
FT
XX
PN
XX
PD
XX
PF
XX
PR
XX
PA
XX
PI
XX
DR
DR
XX
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
XX
SQ
Corynebacterium glutamicum pyruvate carboxylase protein.
Pyruvate carboxylase; anaplerotic enzyme; industrial fermentation;
oxaloacetate; growth; enzyme.
Corynebacterium glutamicum.
Key Location/Qualifiers
Misc-difference 1 /note= "Encoded by GTG"
US6403351-B1.
11-JUN-2002.
03-OCT-2000; 2000US-00677575.
23-DEC-1998; 98US-00220081.
(ARCH) ARCHER-DANIELS MIDLAND CO.
Sinskey AJ, Lessard PA, Willis LB;
WPI; 2002-536037/57.
N-PSDB; AAD42046.
Novel pyruvate carboxylase polypeptide, useful for replenishing
oxaloacetate consumed for biosynthesis during growth, or lysine and
glutamic acid production in industrial fermentation.
Claim 1; Col 29-36; 28pp; English.
The present invention relates to novel pyruvate carboxylase proteins and
polynucleotides encoding such proteins. Sequences of the invention are
important anaplerotic enzymes for replenishing oxaloacetate consumed for
biosynthesis during growth, or lysine and glutamic acid production in
industrial fermentation. The present sequence is C. glutamicum pyruvate
carboxylase protein
Sequence 1140 AA;
Query Match 100.0%; Score 5788; DB 5; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTHTSSTLPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60
Db 1 MSTHTSSTLPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60
QY 61 IGTEGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120
QY 121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
Db 121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
QY 181 VASPELRLKLAATEASREAEAAFGDGA VYVERAVINPQHIEVOILGDHTGEVVHLYERDCS 240
Db 181 VASPELRLKLAATEASREAEAAFGDGA VYVERAVINPQHIEVOILGDHTGEVVHLYERDCS 240
QY 241 LORRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEWN 300
Db 241 LORRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEWN 300
QY 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
QY 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
Db 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420

QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADV 480
Db |||||
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADV 480
Db |||||
QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPGSRDRDLKQLGPAAFARDLREQDALAVTDTT 540
Db |||||
QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPGSRDRDLKQLGPAAFARDLREQDALAVTDTT 540
Db |||||
QY 541 FRDAHQSLLATRVRSFALKPAEAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 600
Db |||||
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFYKAEASSGVDFRIFDALNDVSOQMRPA 660
Db |||||
QY 661 IDAVLETNTAVAEVAMAYSGDLPNEKLYTLDYILKVAEEIVKSGAHILAIDKDMAGLLR 720
Db |||||
QY 721 PAAVTKLVTLRREFDLPVHVHTHTAGGQLATYPAAQAQADAVDASAPLSGTTSQPS 780
Db |||||
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGTVRHEIPGGQL 840
Db |||||
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLRPTKVTTPSSKVVGDLALHLVAGVDPADE 900
Db |||||
QY 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEEQAHLDA 960
Db |||||
QY 961 DDSKERRNSLNRLLFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
Db |||||
QY 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db |||||
QY 1081 GVVTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db |||||
RESULT 4
AAU98053
ID AAU98053 standard; protein; 1140 AA.
XX
AC AAU98053;
XX
DT 27-AUG-2002 (first entry)
XX
DE Corynebacterium wild-type feedback-resistant pyruvate carboxylase enzyme.
XX
KW Feedback-resistant; pyruvate carboxylase; enzyme;
KW aspartic acid feedback inhibition resistant.
XX
OS Corynebacterium glutamicum.
XX
PN WO200231158-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031893.
XX
PR 13-OCT-2000; 2000US-0239913P.
XX

PA (ARCH) ARCHER-DANIELS MIDLAND CO.
XX
PI Hanke PD;
XX
DR WPI; 2002-463267/49.
XX
PT Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT polypeptide, useful for producing amino acids e.g. L-lysine, L-threonine,
PT L-glycine, L-glutamic acid, L-proline and L-methionine and L-isoleucine.
PS Disclosure; Fig 2; 42pp; English.
XX
CC The present invention relates to a new mutated, feedback-resistant
CC pyruvate carboxylase enzyme. The invention is useful for producing an
CC amino acid (e.g. L-Lys, L-Thr, L-Met, L-Phe, L-Glu, L-Arg and L-Pro), by
CC culturing a host cell in a suitable media and separating the amino acid
CC from the medium. The vector of the invention is useful for replacement of
CC a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate
CC carboxylase gene, in a Corynebacterium glutamicum, by replacing a genomic
CC copy of the wild-type pyruvate carboxylase gene with a selectable marker
CC gene through homologous recombination to form a first recombination
CC strain, and replacing the selectable marker gene in the first recombinant
CC strain, with feedback resistant pyruvate carboxylase gene through
CC homologous recombination to form a second recombinant strain, where the
CC homologous recombination in the above steps, occurs between the host cell
CC and the vector. The feedback-resistant pyruvate carboxylase enzyme is
CC resistant to feedback inhibition from aspartic acid. The present amino
CC acid sequence represents the wild-type feedback-resistant pyruvate
CC carboxylase enzyme of the invention
XX
SQ Sequence 1140 AA;

Query Match 100.0%; Score 5788; DB 5; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60
Db 1 MSTHTSSTLPFAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60

QY 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEV 120

QY 121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF 180
Db 121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF 180

QY 181 VASDELRLKLAATESREAEAFDGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240
Db 181 VASDELRLKLAATESREAEAFDGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240

QY 241 LQRRHQKVEIAPAQHLDPFLDRICADAVKFCESIGYQGAGTVEFLVDEKGNHVFIEMN 300
Db 241 LQRRHQKVEIAPAQHLDPFLDRICADAVKFCESIGYQGAGTVEFLVDEKGNHVFIEMN 300

QY 301 PRIQVEHTVTEEVTEVDLVKAQMRLLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

QY 361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
Db 361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420

QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADV 480

QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPGSRDRDLKQLGPAAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPGSRDRDLKQLGPAAFARDLREQDALAVTDTT 540

QY 541 FRDAHQSLLATRVRSFALKPAEAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 600

||||| 541 FRDAHQSLLATVRVSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTALRREFDLPVHVHTHTAGGQLATYFAAAQACADAVDGASAPLSGTTSQPS 780
Db 721 PAAVTKLVTALRREFDLPVHVHTHTAGGQLATYFAAAQACADAVDGASAPLSGTTSQPS 780
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
QY 901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGSEGGKAPLTEVPPEEQAHLDA 960
Db 901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGSEGGKAPLTEVPPEEQAHLDA 960
QY 961 DDSKERRNSLNRLLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
Db 961 DDSKERRNSLNRLLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
QY 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAPFA 1080
Db 1021 LLVRLDAISEPDDKGMRNVVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1081 GVTVTVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
Db 1081 GVTVTVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
RESULT 5
ABU10426
ID ABU10426 standard; protein; 1140 AA.
XX
AC ABU10426;
XX
DT 25-AUG-2003 (first entry)
XX
DE Corynebacterium glutamicum pyruvate carboxylase.
XX
KW Pyruvate carboxylase; gene; anaplerotic enzyme; oxaloacetate;
KW biosynthesis; growth; lysine production; glutamic acid production;
KW industrial fermentation; enzyme.
XX
OS Corynebacterium glutamicum.
XX
PN US2003027305-A1.
XX
PD 06-FEB-2003.
XX
PF 15-JAN-2002; 2002US-00045072.
XX
PR 23-DEC-1998; 98US-00220081.
PR 03-OCT-2000; 2000US-00677575.
XX
PA (ARCH) ARCHER-DANIELS MIDLAND CO.
XX
PI Sinskey AJ, Lessard PA, Willis LB;
XX
DR WPI: 2003-479542/58.
DR N-PSDB; ACA62133.
XX
PT New pyruvate carboxylase from Corynebacterium glutamicum, useful as an

PT anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis
PT during growth, or for lysine or glutamic acid production in industrial
PT fermentations.
XX Claim 1; Page 16-19; 29pp; English.
PS
XX The invention describes a new isolated pyruvate carboxylase polypeptide
CC having an amino acid sequence at least 95% identical to a sequence
CC comprising 1140 amino acids from Corynebacterium glutamicum, or the
CC complete amino acid sequence encoded by the cosmid clone deposited with
CC the American Type Culture Collection. The polypeptide is useful as an
CC anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis
CC during growth. The polypeptide is also useful for lysine or glutamic acid
CC production in industrial fermentations. This is the amino acid sequence
CC of Corynebacterium glutamicum pyruvate carboxylase
XX
SQ Sequence 1140 AA;
Query Match 100.0%; Score 5788; DB 7; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
Db 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSFASEAVR 60
QY 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
QY 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQYPIFVKAVAGGGGRGMRF 180
Db 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQYPIFVKAVAGGGGRGMRF 180
QY 181 VASPELRLKLAATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240
Db 181 VASPELRLKLAATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240
QY 241 LQRRHQKVVEIAPAQHLDPRLDRICADAVKFCRSIGYCGAGTVFVLVDEKGNHVFIEMN 300
Db 241 LQRRHQKVVEIAPAQHLDPRLDRICADAVKFCRSIGYCGAGTVFVLVDEKGNHVFIEMN 300
QY 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
QY 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
Db 361 FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLIQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLIQAPPADDEQGRILDYLDV 480
QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
QY 541 FRDAHQSLLATVRVSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
Db 541 FRDAHQSLLATVRVSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFAVKEAASSGVDIFRIFDALNDVSQMRPA 660
QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTALRREFDLPVHVHTHTAGGQLATYFAAAQACADAVDGASAPLSGTTSQPS 780
Db 721 PAAVTKLVTALRREFDLPVHVHTHTAGGQLATYFAAAQACADAVDGASAPLSGTTSQPS 780

QY 781 LSAIVAAFAHTRRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGGQL 840
DB 781 LSAIVAAFAHTRRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGGQL 840
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
DB 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
QY 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDA 960
DB 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDA 960
QY 961 DDSKERNLSNRLLPFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETILRLPDVRTP 1020
DB 961 DDSKERNLSNRLLPFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETILRLPDVRTP 1020
QY 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
DB 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPATKVEGGDLIVVVS 1140
DB 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPATKVEGGDLIVVVS 1140

RESULT 6
AAB01436
ID AAB01436 standard; protein; 1141 AA.

XX AAB01436;

DT 20-OCT-2000 (first entry)

XX Pyruvate carboxylase of C. glutamicum.

KW Pyruvate carboxylase; expression; amino acid biosynthesis; lysine;
KW glutamic acid; oxaloacetate; fermentation; biosynthesis.

OS Corynebacterium glutamicum.

XX WO200039305-A1.

PN 06-JUL-2000.

XX 23-DEC-1998; 98WO-US027301.

PF 23-DEC-1998; 98WO-US027301.

XX (SINS/) SINSKEY A J.
PA (LESS/) LESSARD P A.
PA (WILL/) WILLIS L B.

XX Sinskey AJ, Lessard PA, Willis LB;

XX WPI; 2000-465746/40.
DR N-PSDB; AAA47533.

XX Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
PT carboxylase useful for industrial fermentation processes comprises a
PT specific nucleotide sequence.

XX Claim 3; Fig 1; Slpp; English.

XX The pyruvate carboxylase of Corynebacterium glutamicum can be used for
CC producing amino acids, preferably lysine and glutamic acid in industrial
CC fermentations and for replenishing oxaloacetate consumed for biosynthesis
CC during growth. By incorporating the pyruvate carboxylase gene in
CC expression vectors levels of expression can be 2 - 20 fold higher than in
CC Corynebacterium glutamicum

XX Sequence 1141 AA;

QY 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080

Query Match 99.9%; Score 5784; DB 3; Length 1141;
Best Local Similarity 99.9%; Pred.No. 0;
Matches 1139; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAPFKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSEAVR 60
DB 2 VSTHTSSTLPAPFKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSFHRSEAVR 61
QY 61 IGTEGSPVKAYLDIDEIIGAAGKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
DB 62 IGTEGSPVKAYLDIDEIIGAAGKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 121
QY 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGOTYPIFVKAVAGGGGRGMR 180
DB 122 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGOTYPIFVKAVAGGGGRGMR 181
QY 181 VASPDRLKCLATEASREAEAFGDGAVYVERAVINPOHIEVQILGDHTGEVWHLYERDCS 240
DB 182 VASPDRLKCLATEASREAEAFGDGAVYVERAVINPOHIEVQILGDHTGEVWHLYERDCS 241
QY 241 LQRRHQKVEIAPAQHLDPDLRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 300
DB 242 LQRRHQKVEIAPAQHLDPDLRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN 301
QY 301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 360
DB 302 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 361
QY 361 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDEETAVARAQAL 420
DB 362 FRPDTGTITAYRSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDEETAVARAQAL 421
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
DB 422 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 481
QY 481 TVNKPVGVRPKDVAAPIIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
DB 482 TVNKPVGVRPKDVAAPIIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 541
QY 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWRLDE 600
DB 542 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPDWRLDE 601
QY 601 LREAMPNVNIQMLLRGNTVGYTPYPDSVCRAFAVKEAASSGVDIFRIFDALNDVQMRPA 660
DB 602 LREAMPNVNIQMLLRGNTVGYTPYPDSVCRAFAVKEAASSGVDIFRIFDALNDVQMRPA 661
QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDDYILKMAEIIKSGAHILAIDMAGLLR 720
DB 662 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDDYILKMAEIIKSGAHILAIDMAGLLR 721
QY 721 PAAVTKLVTALRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTSQPS 780
DB 722 PAAVTKLVTALRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTSQPS 781
QY 781 LSAIVAAFAHTRRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGGQL 840
DB 782 LSAIVAAFAHTRRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTRVYRHEIPGGQL 841
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
DB 842 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 901
QY 901 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDA 960
DB 902 AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDA 961
QY 961 DDSKERNLSNRLLPFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETILRLPDVRTP 1020
DB 962 DDSKERNLSNRLLPFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETILRLPDVRTP 1021

Db 1022 LLVRLDAISEPDDKGRNVVAVNVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1081
QY 1081 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db 1082 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1141

RESULT 7
AAW93971
ID AAW93971 standard; protein; 1140 AA.
XX
AC AAW93971;
XX
DT 30-JUN-1999 (first entry)
XX
DE C. glutamicum pyruvate carboxylase protein.
XX
KW Pyruvate carboxylase; amino acid production; lysine production;
KW threonine production; homoserine production; glutamate production;
KW arginine production; feed additive; condiment; pharmaceutical;
KW fine chemical; ss.
XX
OS Corynebacterium glutamicum.
XX
PN DE19831609-A1.
XX
PD 15-APR-1999.
XX
PF 14-JUL-1998; 98DE-01031609.
XX
PR 04-OCT-1997; 97DE-01043894.
XX
PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
PI Peters-Wendisch P, Eikmanns B, Sahm H;
XX
DR WPI; 1999-245521/21.
XX
PT Increasing microbial production of specific amino acids by increasing
PT activity or expression of pyruvate carboxylase.
XX
PS Claim 15; Page 11-15; 18pp; German.
XX
CC This invention describes the isolation of a pyruvate carboxylase from
CC Corynebacterium glutamicum which is used in a novel method for production
CC of lysine, threonine, homoserine, glutamate and/or arginine, variously
CC useful as feed additives, condiments, pharmaceuticals and intermediates
CC for fine chemicals. Increasing pyruvate carboxylase activity increases
CC the yield of microbial production of amino acids of the asparate and/or
CC glutamate families, e.g. about 50% more lysine, 40% more threonine and
CC 150% more homoserine are secreted into the culture medium
XX
SQ Sequence 1140 AA;
Query Match 99.9%; Score 5782; DB 2; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAPKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVR 60
Db 1 MSTHTSSTLPAPKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVR 60

QY 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

QY 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
Db 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

QY 181 VASPELRLKLAATESREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240

Db 181 VASPELRLKLAATESREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
QY 241 LQRHOKVVEIAPAOHLDPPELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300
Db 241 LQRHOKVVEIAPAOHLDPPELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300
QY 301 PRIQVEHTVTEEVTEVDLVKAQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 360
QY 361 FRPDTGITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db 361 FRPDTGITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
QY 481 TVNKPVGVRPKDVAAAPIDKLPNIDPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAAPIDKLPNIDPLPRGSRDRLKQLGPAAFARDLREQDALAVTDTT 540
QY 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
Db 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYPSVCRAFAVKEAASSGVDIIFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRGRNTVGYTPYPSVCRAFAVKEAASSGVDIIFRIFDALNDVSQMRPA 660
QY 661 IDAVLETNTAVAEVAMAYSGLSDPNNEKLYTLDYLLKMAEEIVKGAHILAIKDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGLSDPNNEKLYTLDYLLKMAEEIVKGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVTLRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDCASAPLSGTTSQPS 780
Db 721 PAAVTKLVTLRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDCASAPLSGTTSQPS 780
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADF 900
QY 901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHLLDA 960
Db 901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPPEEQAHLLDA 960
QY 961 DDSKERRNSLNRLFPKPTEEFLEHRRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db 961 DDSKERRNSLNRLFPKPTEEFLEHRRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
QY 1021 LLVRLDAISEPDDKGRNVVAVNVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db 1021 LLVRLDAISEPDDKGRNVVAVNVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1081 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140

RESULT 8
AAG93249
ID AAG93249 standard; protein; 1140 AA.
XX
AC AAG93249;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment mutant P458S.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; mutant; mutein.
KW

OS Corynebacterium glutamicum.
OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 458
FT /note= "wild-type Pro substituted by Ser"

XX PN EP1108790-A2.

XX PD 20-JUN-2001.

XX PF 18-DEC-2000; 2000EP-00127688.

XX PR 16-DEC-1999; 95JP-00377484.

XX PR 07-APR-2000; 2000JP-00159162.

XX PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX DR WPI; 2001-376931/40.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.

PS Claim 43; Page; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a mutant protein described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

XX Sequence 1140 AA;

Query Match 99.9%; Score 5780; DB 4; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAPFKKILVANRGEIAVRAALETGAATVAIYPREDRGSFHRSPASEAVR 60

Db 1 MSTHTSSTLPAPFKKILVANRGEIAVRAALETGAATVAIYPREDRGSFHRSPASEAVR 60

QY 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

Db 61 IGTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120

QY 121 LDLTGDKSRVATAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF 180

Db 121 LDLTGDKSRVATAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF 180

QY 181 VASPELRLKATEASREAEAFDGDGAVYVERAVINPOHIEVQILGDHTGEVHLYERDCS 240

Db 181 VASPELRLKATEASREAEAFDGDGAVYVERAVINPOHIEVQILGDHTGEVHLYERDCS 240

QY 241 LQRHQKVVEIAPAQHLDPPELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300

Db 241 LQRHQKVVEIAPAQHLDPPELDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300

QY 301 PRIQVEHTVTEEVTEVDLVKQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
QY 361 FRPDTGTITAYRSPGGAGVRLDGAAGLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db 361 FRPDTGTITAYRSPGGAGVRLDGAAGLGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
QY 421 AEFTVSGVATNIGFLRALRREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALRREEDFTSKRIATGFIADHSHLLQAPPADDEQGRILDYLDV 480
QY 481 TVNKPBGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
Db 481 TVNKPBGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540
QY 541 FRDAHOSLLATRVRSFALKPAABEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 600
Db 541 FRDAHOSLLATRVRSFALKPAABEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDE 600
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFFVKEAASSGVDFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFFVKEAASSGVDFRIFDALNDVSQMRPA 660
QY 661 IDAVLETNTAVAEVAMAYSGDSDPNEKLYTLDYLLKVAEEIVKSGAHILAIDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGDSDPNEKLYTLDYLLKVAEEIVKSGAHILAIDMAGLLR 720
QY 721 PAAVTKLVTALRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTTSPS 780
Db 721 PAAVTKLVTALRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTTSPS 780
QY 781 LSAIVAAPFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGSQL 840
Db 781 LSAIVAAPFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGSQL 840
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVDLALHLVGAGVDPADF 900
QY 901 AADPOKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEEQAHLDA 960
Db 901 AADPOKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEEQAHLDA 960
QY 961 DDSKERRNSLNRLLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
Db 961 DDSKERRNSLNRLLFPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020
QY 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db 1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRDRSVESVTATAEKADSSNKGHVAAPFA 1080
QY 1081 GVVTVTVAEAGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db 1081 GVVTVTVAEAGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
RESULT 9
AAU98052
ID AAU98052 standard; protein; 1140 AA.
XX AC AAU98052;
XX AC AAU98052;
XX AC AAU98052;
DT 27-AUG-2002 (first entry)
XX Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
DE Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
XX Feedback-resistant; pyruvate carboxylase; enzyme;
KW aspartic acid feedback inhibition resistant; mutant; mutein.
XX
OS Corynebacterium glutamicum.
OS Synthetic.
XX

FH Key Location/Qualifiers
FT Misc-difference 1 /note= "Wild-type Met substituted by Val"
FT Misc-difference 153 /note= "Wild-type Glu substituted by Asp"
FT Misc-difference 182 /note= "Wild-type Ala substituted by Ser"
FT Misc-difference 206 /note= "Wild-type Ala substituted by Ser"
FT Misc-difference 227 /note= "Wild-type His substituted by Arg"
FT Misc-difference 455 /note= "Wild-type Ala substituted by Gly"
FT Region 1110. .1122 /note= "Specifically claimed in claim 18"
FT Misc-difference 1116 /note= "Wild-type Asp substituted by Glu"
FT
XX
PN WO200231158-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031893.
XX
PR 13-OCT-2000; 2000US-0239913P.
XX
PA (ARCH) ARCHER-DANIELS MIDLAND CO.
XX
PI Hanke PD;
XX
XX WPI; 2002-463267/49.
DR
XX
XX Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT polypeptide, useful for producing amino acids e.g. L-lysine, L-threonine,
PT L-glycine, L-glutamic acid, L-proline and L-methionine and L-isoleucine.
XX
PS Claim 1; Page; 42pp; English.
XX
CC The present invention relates to a new mutated, feedback-resistant
CC pyruvate carboxylase enzyme. The invention is useful for producing an
CC amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro), by
CC culturing a host cell in a suitable media and separating the amino acid
CC from the medium. The vector of the invention is useful for replacement of
CC a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate
CC carboxylase gene, in a Corynebacterium glutamicum, by replacing a genomic
CC copy of the wild-type pyruvate carboxylase gene with a selectable marker
CC gene through homologous recombination to form a first recombinant
CC strain, and replacing the selectable marker gene in the first recombinant
CC strain, with feedback resistant pyruvate carboxylase gene through
CC homologous recombination to form a second recombinant strain, where the
CC homologous recombination in the above steps, occurs between the host cell
CC and the vector. The feedback-resistant pyruvate carboxylase enzyme is
CC resistant to feedback inhibition from aspartic acid. The present amino
CC acid sequence represents the mutant feedback-resistant pyruvate
CC carboxylase enzyme of the invention. Note: The present sequence is not
CC shown in the specification but is derived from the wild-type feedback-
CC resistant pyruvate carboxylase enzyme (AAU98053) given in figure 2 of the
CC specification
XX
SQ Sequence 1140 AA;

Query Match 99.5%; Score 5759; DB 5; Length 1140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAFKKILVANRGEIAVRAFAALETGAATVAIYPREDGSGFHRSPASEAVR 60
Db :|||||
QY 1 VSTHTSSTLPAFKKILVANRGEIAVRAFAALETGAATVAIYPREDGSGFHRSPASEAVR 60
Db :|||||
QY 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
Db :|||||
QY 61 IGTEGSPVKAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV 120
Db :|||||

QY 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSABGOTYPIFVKAVAGGGRGMRF 180
Db :|||||
QY 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDIVKSABGOTYPIFVKAVAGGGRGMRF 180
Db :|||||
QY 181 VASPDRLKRLATEASREAEAAFGDGAVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240
Db :|||||
QY 181 VSSPDRLKRLATEASREAEAAFGDGSVYVERAVINPQHIEVQILGDRTEVHLYERDCS 240
Db :|||||
QY 241 LORRHQKVVEIAPAQHLDPRLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300
Db :|||||
QY 301 PRIQVEHTVTEEVTEVDLVKAQVRLAAGATLKELGLTDQIKITHGAALQCRITTEDPNNG 360
Db :|||||
QY 301 PRIQVEHTVTEEVTEVDLVKAQVRLAAGATLKELGLTDQIKITHGAALQCRITTEDPNNG 360
Db :|||||
QY 361 FRPDTGITAYRSPGGAGVRLDGAQAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db :|||||
QY 361 FRPDTGITAYRSPGGAGVRLDGAQAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db :|||||
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480
Db :|||||
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADV 480
Db :|||||
QY 481 TVNKPVGVRPKDVAAPIDKLPNIDKLPGRSRDLKQLGPAAFARDLREODALAVTDTT 540
Db :|||||
QY 481 TVNKPVGVRPKDVAAPIDKLPNIDKLPGRSRDLKQLGPAAFARDLREODALAVTDTT 540
Db :|||||
QY 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEANGGATYDVAMRFLFEDPWRDLDE 600
Db :|||||
QY 541 FRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEANGGATYDVAMRFLFEDPWRDLDE 600
Db :|||||
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYDSDVCRAFVKEAASSGVDFRIFDALNDVSQMRPA 660
Db :|||||
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYDSDVCRAFVKEAASSGVDFRIFDALNDVSQMRPA 660
Db :|||||
QY 661 IDAVLETNTAVAEVAMAYSGDLSDENKLYTLDYLLKMAEIEVKSAGAHILAIDMAGLLR 720
Db :|||||
QY 661 IDAVLETNTAVAEVAMAYSGDLSDENKLYTLDYLLKMAEIEVKSAGAHILAIDMAGLLR 720
Db :|||||
QY 721 PAAVTKLVTLARRREFDLFVHVHTHTAGGQLATYFAAAQAGADAVDASAPLSGTTSQPS 780
Db :|||||
QY 721 PAAVTKLVTLARRREFDLFVHVHTHTAGGQLATYFAAAQAGADAVDASAPLSGTTSQPS 780
Db :|||||
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db :|||||
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db :|||||
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
Db :|||||
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
Db :|||||
QY 901 AADPQKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTPVEPEEQAHLLDA 960
Db :|||||
QY 901 AADPQKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTPVEPEEQAHLLDA 960
Db :|||||
QY 961 DDSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db :|||||
QY 961 DDSKERRNSLNRLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRT 1020
Db :|||||
QY 1021 LLVRLDAISEPDDKGMNVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db :|||||
QY 1021 LLVRLDAISEPDDKGMNVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db :|||||
QY 1081 GWVTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
Db :|||||
QY 1081 GWVTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
Db :|||||

RESULT 10
AAU98050
ID AAU98050 standard; protein; 1157 AA.
XX

AC AAU98050;
XX
DT 27-AUG-2002 (first entry)
XX
DE Corynebacterium feedback-resistant pyruvate carboxylase enzyme.
XX
KW Feedback-resistant; pyruvate carboxylase; enzyme;
KW aspartic acid feedback inhibition resistant.
XX
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT Region 1..18
FT /note= "Specifically claimed in claim 18"
FT Region 164..176
FT /note= "Specifically claimed in claim 18"
FT Region 193..205
FT /note= "Specifically claimed in claim 18"
FT Region 217..229
FT /note= "Specifically claimed in claim 18"
FT Region 238..250
FT /note= "Specifically claimed in claim 18"
FT Region 466..478
FT /note= "Specifically claimed in claim 18"

WO200231158-A2.
18-APR-2002.
12-OCT-2001; 2001WO-US031893.
13-OCT-2000; 2000US-0239913P.
(ARCH) ARCHER-DANIELS MIDLAND CO.
Hanke PD;
WPI; 2002-463267/49.
N-PSDB; ABK52832.
Novel mutated, feedback resistant pyruvate carboxylase enzyme
polypeptide, useful for producing amino acids e.g. L-lysine, L-threonine,
L-glycine, L-glutamic acid, L-proline and L-methionine and L-isoleucine.
Claim 17; Fig 1; 42pp; English.

The present invention relates to a new mutated, feedback-resistant
pyruvate carboxylase enzyme. The invention is useful for producing an
amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro), by
culturing a host cell in a suitable media and separating the amino acid
from the medium. The vector of the invention is useful for replacement of
a wild-type pyruvate carboxylase gene, with a feedback resistant pyruvate
carboxylase gene, in a Corynebacterium glutamicum, by replacing a genomic
copy of the wild-type pyruvate carboxylase gene with a selectable marker
gene through homologous recombination to form a first recombinant
strain, and replacing the selectable marker gene in the first recombinant
strain, with feedback resistant pyruvate carboxylase gene through
homologous recombination to form a second recombinant strain, where the
homologous recombination in the above steps, occurs between the host cell
and the vector. The feedback-resistant pyruvate carboxylase enzyme is
resistant to feedback inhibition from aspartic acid. The present amino
acid sequence represents the feedback-resistant pyruvate carboxylase
enzyme of the invention

Sequence 1157 AA;
Query Match 99.5%; Score 5759; DB 5; Length 1157;
Best local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSTHTSSTLPAPKKILVANRGEIAVRAPAALETGAATVAIYPREDRGSFHRSPASEAVR 60
DB 18 VSTHTSSTLPAPKKILVANRGEIAVRAPAALETGAATVAIYPREDRGSFHRSPASEAVR 77

QY 61 IGTEGSPVKAYLDIDEIIGAAGKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEV 120
DB 78 IGTEGSPVKAYLDIDEIIGAAGKVKADAIYPGYFLSENAQLARECAENGITFIGPTPEV 137
QY 121 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMR 180
DB 138 LDLTGDKSRAVTAAKKAGLPVLAESTPSKNIDIVKSAEGQTYPIFVKAVAGGGRGMR 197
QY 181 VASPDRLKLAATESREAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
DB 198 VSSPDRLKLAATESREAAFGDGSVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 257
QY 241 LQRRHOKVWEIAPAQHLDPDLDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIE 300
DB 258 LQRRHOKVWEIAPAQHLDPDLDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIE 317
QY 301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALQCRITTEDPN 360
DB 318 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALQCRITTEDPN 377
QY 361 FRPDTGITAYRSPGGAGVRLDGAAGLGEITAHFDSMLVKMTCRGSDFETAVARAQR 420
DB 378 FRPDTGITAYRSPGGAGVRLDGAAGLGEITAHFDSMLVKMTCRGSDFETAVARAQR 437
QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDY 480
DB 438 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDY 497
QY 481 TVNKPCHGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAV 540
DB 498 TVNKPCHGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAV 557
QY 541 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEANGGATYDVAMRFLFEDP 600
DB 558 FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEANGGATYDVAMRFLFEDP 617
QY 601 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFVKEAASSGVDFRIFDALNDV 660
DB 618 LREAMPNVNIQMLLRGRNTVGYTPYDPSVCRAFVKEAASSGVDFRIFDALNDV 677
QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDDYILKMAEEIVKSGAHILA 720
DB 678 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDDYILKMAEEIVKSGAHILA 737
QY 721 PAAVTKLVTLRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSG 780
DB 738 PAAVTKLVTLRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSG 797
QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVYRHE 840
DB 798 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVYRHE 857
QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTSSKVVGDLALHLVGAG 900
DB 858 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTSSKVVGDLALHLVGAG 917
QY 901 AADPQKYDIPDSVIAFLRGLGNPFGGWPEPLRTRALEGRSEKAPLTEVP 960
DB 918 AADPQKYDIPDSVIAFLRGLGNPFGGWPEPLRTRALEGRSEKAPLTEVP 977
QY 961 DDSKERNLSNRLLPFKPTTEEFLEHRRRRFGNTSALDDREFFYGLVEGRETL 1020
DB 978 DDSKERNLSNRLLPFKPTTEEFLEHRRRRFGNTSALDDREFFYGLVEGRETL 1037
QY 1021 LLVRLDAISEPDDKGMNRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKG 1080
DB 1038 LLVRLDAISEPDDKGMNRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKG 1097
QY 1081 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGG 1140
DB 1098 GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGG 1157

RESULT 11	
AAB83180	
ID	AAB83180 standard; protein; 1139 AA.
XX	
AC	AAB83180;
XX	
DT	09-JUL-2001 (first entry)
XX	
DE	Corynebacterium thermoaminogenes pc protein.
XX	
KW	Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW	thermotolerant; aceA; accBC; dtsR1; dtsR2; pfk; scrB; gluABCD; pdhA; pc;
KW	ppc; acn; icd; lpd; odhA.
XX	
OS	Corynebacterium thermoaminogenes.
XX	
PN	WO200125447-A1.
XX	
PD	12-APR-2001.
XX	
FF	04-OCT-2000; 2000WO-JP006913.
XX	
PR	04-OCT-1999; 99JP-00282716.
PR	01-NOV-1999; 99JP-00311147.
PR	21-APR-2000; 2000JP-00120687.
XX	
PA	(AJIN) AJINOMOTO CO INC.
XX	
PI	Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K;
PI	Kimura E, Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
PI	Sugimoto S;
XX	
DR	WPI; 2001-300170/31.
XX	
PT	N-PSDB; AAF87437.
XX	
PS	Proteins and their DNA useful for microbial production of L-amino acids.
XX	
CC	Claim 9; Page 132-135; 215pp; Japanese.
CC	The present sequence is provided in a specification relating to genes encoding thermophilic amino acid biosynthesis system enzymes of the thermotolerant bacterium Corynebacterium thermoaminogenes. The novel proteins retain at least 30% isocitrate ligase activity after heating at 50oC for 5 minutes. DNA fragments encoding the enzymes were isolated from a Corynebacterium thermoaminogenes chromosomal DNA plasmid library by PCR. The DNA may be used for developing strains of amino acid producing microorganisms
XX	
SQ	Sequence 1139 AA;
Query Match	91.6%; Score 5302.5; DB 4; Length 1139;
Best Local Similarity	91.0%; Pred. No. 0;
Matches 1036;	Conservative 43; Mismatches 58; Indels 1; Gaps 1;
QY	3 THTSSTLPFAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVRIG 62
Db	3 TTPSTLPFAFKKILVANRGEIAVRAFAAYETGAATVAIYPREDRGSFHRSPASEAVRIG 62
QY	63 TEGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLD 122
Db	63 TEGSPVKAYLDIDEIINAACKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEVLD 122
QY	123 LTGDKSRVATAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMFEVA 182
Db	123 LTGDKSKAVSAACKAGLPVLAESTPTDIDEIVKSAEGQTYPIFVKAVAGGGGRGMFEVE 182
QY	183 SPDELKRLATEASREAEAFGDGAVYVERAVINPQHIEVQILGDHGTGEVVLHYERDCSLQ 242
Db	183 KPEDLRELAREASREAEAFGDGSGVYVERAVINPQHIEVQILGDHGTGDVHLHYERDCSLQ 242
QY	243 RRHOKVVEIAPACHLDPDLRDRICADAVKFRSIGYQGAGTVEFLVDEKGNHVFIEMNPR 302
Db	243 RRHOKVVEIAPACHLDPDLRDRICADAVKFKSIGYQGAGTVEFLVDEAGNHVFIEMNPR 302
QY	303 IQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGITQDKIKTHGAALQCRITTEDPNNGFR 362
Db	303 IQVEHTVTEEVTSVDLVKAQMHLAAGATLKELGITQDKIKTHGAALQCRITTEDPSNNFR 362
QY	363 PDTGITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQALAE 422
Db	363 PDTGVITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFETAVARAQALAE 422
QY	423 FTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYADTV 482
Db	423 FNVSGVATNIGFLRALLREEDFTKRRIDTGFICSHQHLQAPPADDEQGRILEYADTV 482
QY	483 NKPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVDTTFR 542
Db	483 NKPHGERP-ETARPIEKLPEVENIPLPRGSRDLKQLGPEGFARDLREQDALAVDTTFR 541
QY	543 DAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLLDEL 602
Db	542 DAHQSLLATRVRSFALTFAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDEL 601
QY	603 EAMPNVNIQMLLRGNTVGYTPYDPSVCRAFYKAAASSGVDFIRIFDALNDVSQMRPAID 662
Db	602 EAMPNVNIQMLLRGNTVGYTPYDPSVCRAFYQVQEAASKGVDFIRIFDALNDISQMRPAID 661
QY	663 AVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYVYKMAEEIVKSGAHILAIKDMAGLLRPA 722
Db	662 AVLETGTSVAEVAMAYSGDLSNPGEKLYTLDYVYLNLAEQIVDSGAHILAIKDMAGLLRRA 721
QY	723 AVTKLVALTALRRREFDLPVHVHTHTDAGGQLATYFAAAQAGADAVDGASAPLSGTTSPSLS 782
Db	722 AAPKLVALTALRRREFDLPVHVHTHTDAGGQLATYLAANAAGADAVDAASAPLSGTTSPSMS 781
QY	783 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN 842
Db	782 ALVAAFAHTRRDTGLNLQAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN 841
QY	843 LRAQATALGLADRFELIEDNYAAVNEMLRPTKVTTPSSKVVGDLLHLVAGVDPADFAA 902
Db	842 LRAQAVAGLADRFELIEDYAAVNEMLRPTKVTTPSSKVVGDLLHLVAGVSPEDFAA 901
QY	903 DPQYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDADD 962
Db	902 DPQYDIPDSVIAFLRGELGTPPGGWPEPLRTRALEGRSQKAPLAEIPAEEQAHLDSD 961
QY	963 SKERRNSLRLFFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLL 1022
Db	962 SAERRGTNRLFFKPTEEFLEHRRRFGNTSALDDREFFYGLKEGREELIRLTGVSTPMV 1021
QY	1023 VRLDAISEPDDKGMNVVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFAGV 1082
Db	1022 VRLDAVSEPDCKGMNVVANVNGQIRPIKVRDRSVESVTATAEKADATNKGHVAAPFAGV 1081
QY	1083 VTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db	1082 VTVTVAEGDEIKAGDAVAIEAMKMEATITAPVDGVIDRVVVPAAATKVEGGDLIVVVS 1139
RESULT 12	
ABU25961	
ID	ABU25961 standard; protein; 1141 AA.
XX	
AC	ABU25961;
XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Protein encoded by Prokaryotic essential gene #11488.
XX	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
OS	Corynebacterium diphtheriae.
XX	

PN WO200277183-A2.
XX 03-OCT-2002.
PD 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA29831.
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 53885; 1766pp; English.
PS The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1141 AA;
SQ Query Match 80.3%; Score 4647.5; DB 6; Length 1141;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 917; Conservative 81; Mismatches 137; Indels 3; Gaps 2;
QY 2 STHSTSLPAFKKILVANRGEIAVRAALETGAATVAVIYPREDRGSFHRSFASEAVRI 61
DB 6 SERGSTVNPLSKILVANRGEIAVRAAFETGAATVAVIYPNEDRNSFHRSFASEAVLI 65
QY 62 GTEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFISENAQLARECAENGITFGPTPEVL 121
DB 66 GEGSAVKAYLDIDEIIRAQKQTGADAIYPGYGFISENAQLARECAENGITFGPPPSVL 125
QY 122 DLTGDKSRVATAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRFV 181

Db 126 ELTGDKAAAATAAREAGLPTLTETETATDDPKLAELTKGQTYPLFVKAVAGGGGRMRFV 185
QY 182 ASPDELKRLATEASREAEAFDGAAYVERAVINPQHIEVQILGDHTGEVVLHVERDCSL 241
Db 186 DAPENLEKLAEEASREAAAFDGRVYAEAVINPQHIEVQILGDSAGNIHLYERDCSL 245
QY 242 QRRHQKVVEIAPAQHLDPEDRLICADAVKFCRSIGYQAGTVEFLVDEKGNHVFEMNP 301
Db 246 QRRHQKVVEIAPAQHLAPGLREKICADAVAFARHIGYQAGTVEFLVDEKGNHVFEMNP 305
QY 302 RIQVEHTVTEEVTEVDLVKAQMLAAAGATLKELGLTDQIKTHGAALQCRITTEPNNGF 361
Db 306 RIQVEHTVTEEVTVQDLVKSQIMIASGATLEDGLRQEDIHTEGAALQCRITTEPNNGF 365
QY 362 RPDGTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVQMTCRGSDFETAVARAQALA 421
Db 366 RPDGTITAYRSPGGAGVRLDGAALGGEITPNFDSMLVQMTCRGADFATAVARAQRALA 425
QY 422 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEOQRILDLADVT 481
Db 426 EFTVSGVATNIGFLRALLREEDFQHKRIATGFIADHPHLLQAPPADDEPGRILNYLADVT 485
QY 482 VNKPHGVPRPKDVAAPIDKLPNIKDLPLPRGSRDRLLQGLPEGFARALRKQDALAVTDTTF 541
Db 486 VNKPHGLRPA-VINPVEKLPABEIKGELPRGSRDRLLQGLPEGFARALRKQDALAVTDTTF 544
QY 542 RDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAGGATYDVAMRFLFEDPWRDLDEL 601
Db 545 RDAHQSLLATRVRSNTLIDAAHVAKLTPPELLSVEAGGATYDVAMRFLHEDPWERLDHL 604
QY 602 REAMPNVNIQMLLRNTVGYTPYDPSVCRAFVKEAASGVDFIRIFDALNDVSOVRPAI 661
Db 605 REAMPNVNIQMLLRNTVGYTPYDPSVCRAFVKEAASGVDFIRIFDALNDVSOVRPAI 664
QY 662 DAVLETNTTAEVAMAYSGDLTPNEKLYTLDDYVLMABEIVKSGAHILAIDMAGLMLRP 721
Db 665 DAVLETNTTAEVAMAYSGDLTPNEKLYTLDDYVLMABEIVKSGAHILAIDMAGLMLRP 724
QY 722 AAVTKLVTLRREFDLPVHVTHTDTAGGQATYFAAAQAGADAVDGASAPLSGTSQPSL 781
Db 725 AAVTKLVTLRREFDLPVHVTHTDTAGGQATYFAAAQAGADAVDGASAPLSGTSQPSL 784
QY 782 SAIVAAPAFHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLS 841
Db 785 SAIVAAPAFHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLS 844
QY 842 NLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKWVGDALHLVAGVDPADFA 901
Db 845 NLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKWVGDALHLVAGVDPADFA 904
QY 902 ADPOKYDIPDSVIAFLRGEIENPPGGWPEPLRTRALEGRSEKAPLTVPEPEQAHLAD 961
Db 905 ADPOKYDIPDSVIAFLRGEIENPPGGWPEPLRTRALEGRSEKAPLTVPEPEQAHLAD 962
QY 962 DSKERNLSNRLLPKPTPEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPL 1021
Db 963 DRATVRPTLDRLLFPKPAEFAEHRRRQFGDTTKLGDAEFLYGLKEGKETVIRTADSSVPM 1022
QY 1022 LVRLDAISEPDDKGMNVTVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAPFAG 1081
Db 1023 LVRLDAISEPDDKGMNVTVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAPFAG 1082
QY 1082 VVTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVPAATKVVEGDLIVV 1139
Db 1083 VVTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVPAATKVVEGDLIVV 1140

RESULT 13
ABU34026
ID ABU34026 standard; protein; 1127 AA.
XX
AC ABU34026;


```
Db      | : | | : | | | | | | | | : | | : | | | | : |
       1077 GRQVEAGQTIIATIEANKWEAAVTSFGKVARIASVRTAQVEGGDLLMVI 1126
```

RESULT 14

ABU36886
ID ABU36886 standard; protein; 1127 AA.
XX
AC ABU36886;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #22413.

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS *Mycobacterium tuberculosis*.

PN WO200277183-A2.

03-OCT-2002
PD
PD

21-MAR-2002: 2002WO-US0009107

27-MAR-2001: 2001US-00815242

PK ZI-NAK-2001; 2001US-00815242.
PR 06-SEP-2001: 2001US-00848893

PK 08-SEP-2001; 2001US-00348993.
PR 25-OCT-2001; 2001US-0342833B

PK Z3-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00073851

FR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362688P

XX
PA
(ELITE) ELITE PHARM INC

P I Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX

DR WPI; 2003-029926/02.
DR N-PSDB; ACA40756.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 64810; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained

Search completed: March 24, 2004, 22:42:35
Job time : 69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:39:57 ; Search time 57 Seconds
(without alignments)
6310.364 Million cell updates/sec

Title: US-10-045-072-2
Perfect score: 5788
Sequence: 1 MSTHTSSSTLPFAFKILVANR.....RVVVPAAKVEGGDLIVVVS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5788	100.0	1140	16	O54587
2	5759	99.5	1140	2	Q8L2G4
3	5317.5	91.9	1168	16	Q8FRQ0
4	5305.5	91.7	1139	2	Q8RQL2
5	3799	65.6	1124	16	Q9RK64
6	3783.5	65.4	1127	2	Q9F843
7	3712.5	64.1	1127	16	P95127
8	3712.5	64.1	1127	16	Q7TXJ1
9	3483	60.2	1124	2	Q50450
10	3248	56.1	1131	16	Q83HF3
11	3241	56.0	1131	16	Q83FSS
12	2552	44.1	1178	13	Q8JHF6
13	2540.5	43.9	1178	6	Q7YS28
14	2528	43.7	1148	16	Q819M9
15	2526	43.6	1148	16	Q81MT6
16	2525.5	43.6	1150	16	Q9K9M0

17	2525.5	43.6	1178	6	Q866R1
18	2524	43.6	1148	16	Q9KWU4
19	2519.5	43.5	1180	13	Q9DDT1
20	2510	43.4	1166	16	Q7UES1
21	2509.5	43.4	1192	3	Q9HES8
22	2505	43.3	1147	2	P94448
23	2495	43.1	1167	16	Q7WPS4
24	2493	43.1	1146	16	Q8Y846
25	2491.5	43.0	1144	16	Q97FR7
26	2490	43.0	1146	16	Q92CW1
27	2489	43.0	1150	16	Q8NX69
28	2488	43.0	1150	16	Q99UY8
29	2487.5	43.0	1193	3	Q93918
30	2477.5	42.8	1158	16	Q8FYT0
31	2476	42.8	1167	16	Q7WBT0
32	2473.5	42.7	1174	16	Q8UBX3
33	2470.5	42.7	1152	16	Q98F27
34	2464.5	42.6	1158	16	Q8YJ20
35	2464	42.6	1153	16	Q8CPM3
36	2464	42.6	1175	5	O17732
37	2463.5	42.6	1142	16	Q831P4
38	2462.5	42.5	1179	13	Q8AYN3
39	2443	42.2	1185	3	Q9UUE1
40	2441	42.2	1181	5	Q9XZ00
41	2437.5	42.1	1175	3	Q8X1T3
42	2435	42.1	1152	16	Q92L13
43	2435	42.1	1185	3	P78822
44	2435	42.1	1197	5	Q8MKW5
45	2419	41.8	1195	5	Q16921

ALIGNMENTS

RESULT 1

O54587	ID	O54587	PRELIMINARY;	PRT;	1140 AA.
AC	O54587;				
DT	01-JUN-1998 (Tremblrel. 06, Created)				
DT	01-JUN-1998 (Tremblrel. 06, Last sequence update)				
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)				
DE	Pyruvate carboxylase (EC 6.4.1.1).				
GN	pyc OR CGL0689.				
OS	Corynebacterium glutamicum (Brevibacterium flavum).				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
OX	NCBI_TaxID=1718;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=21253;				
RA	Koffas M.A.G., Ramamoorthi R., Pine W.A., Sinskey A.J.,				
RA	Stephanopoulos G.;				
RL	Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 13032;				
RA	Peters-Wendisch P.G., Kreutzer C., Kalinowski J., Patek M., Sahn H.,				
RA	Eikmanns B.J.;				
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;				
RA	Nakagawa S.;				
RT	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";				
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.				
CC	-1- COFACTOR: BIOTIN (BY SIMILARITY).				
DR	EMBL; AF038548; AAB92588.1; -.				
DR	EMBL; Y09548; CAA70739.1; -.				
DR	EMBL; AP005276; BAB98082.1; -.				
DR	HSSP; P24182; IBNC.				
DR	GO; GO:0005737; C:cytoplasm; IEA.				
DR	GO; GO:0005524; F:ATP binding; IEA.				
DR	GO; GO:0009374; F:biotin binding; IEA.				

Q866r1	bos taurus
Q9kwu4	bacillus su
Q9ddt1	brachydanio
Q7ues1	rhodopirell
Q9hes8	aspergillus
P94448	bacillus st
Q7wps4	bordetella
Q8y846	listeria mo
Q97fr7	clostridium
Q92cw1	listeria in
Q8nx69	staphylococ
Q99uy8	staphylococ
Q93918	aspergillus
Q8fyt0	brucella su
Q7wbt0	bordetella
Q8ubx3	agrobacteri
Q98f27	rhizobium l
Q8yj20	brucella me
Q8cpm3	staphylococ
O17732	caenorhabdi
Q83lp4	enterococcu
Q8ayn3	pagrus majo
Q9uel1	schizosacch
Q9xz00	drosophila
Q8xlt3	pichia angu
Q92l13	rhizobium m
P78822	schizosacch
Q8mkw5	drosophila
Q16921	aedes aegyp

DR	GO; GO:0016874; F:ligase activity; IEA.	
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.	
DR	GO; GO:0006094; P:gluconeogenesis; IEA.	
DR	GO; GO:0008152; P:metabolism; IEA.	
DR	InterPro; IPR001882; Biotin_BS.	
DR	InterPro; IPR005482; Biotin_carb_C.	
DR	InterPro; IPR000089; Biotin_lipoYL.	
DR	InterPro; IPR005479; Cbase_L_D2.	
DR	InterPro; IPR005481; Cbase_L_N.	
DR	InterPro; IPR000891; HMGL-like.	
DR	InterPro; IPR003379; PYC_OADA.	
DR	InterPro; IPR005930; Pyruv_carbox.	
DR	Pfam; PF02785; Biotin_carb_C; 1.	
DR	Pfam; PF00364; biotin_lipoYL; 1.	
DR	Pfam; PF00289; CPSase_L_chain; 1.	
DR	Pfam; PF02786; CPSase_L_D2; 1.	
DR	Pfam; PF00682; HMGL-like; 1.	
DR	Pfam; PF02436; PYC_OADA; 1.	
DR	TIGRFAMs; TIGR01235; pyruv_carbox; 1.	
DR	PROSITE; PS00188; BIOTIN; 1.	
DR	PROSITE; PS00867; CPSASE_2; 1.	
KW	Biotin; Ligase; Pyruvate; Complete proteome.	
SQ	SEQUENCE 1140 AA; 123102 MW; 2A6D4B4ED2FEB531 CRC64;	
Query Match 100.0%; Score 5788; DB 16; Length 1140;		
Best Local Similarity 100.0%; Pred. No. 1.4e-299;		
Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MSTHTSSILPAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60	
Db	1 MSTHTSSILPAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVR 60	
QY	61 IGTEGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFTGPTPEV 120	
Db	61 IGTEGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFTGPTPEV 120	
QY	121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAESQTYPIFVKAVAGGGRGMRF 180	
Db	121 LDLTGDKSRAVTAACKAGLPVLAESTPSKNIDEIVKSAESQTYPIFVKAVAGGGRGMRF 180	
QY	181 VASPELEKLEATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240	
Db	181 VASPELEKLEATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVHLYERDCS 240	
QY	241 LQRHQQVVEIAPAQHLDPQLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300	
Db	241 LQRHQQVVEIAPAQHLDPQLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN 300	
QY	301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 360	
Db	301 PRIQVEHTVTEEVTEVDLVKAQMLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNG 360	
QY	361 FRPDTGTTIAYRSPGGAGVRLDGAALQGGEITAHFDSMLVQWTCRGSDFFETAVARAQAL 420	
Db	361 FRPDTGTTIAYRSPGGAGVRLDGAALQGGEITAHFDSMLVQWTCRGSDFFETAVARAQAL 420	
QY	421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480	
Db	421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV 480	
QY	481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540	
Db	481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAVTDTT 540	
QY	541 FRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEANGGATYDVAMRFLFEDPDWRLDE 600	
Db	541 FRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEANGGATYDVAMRFLFEDPDWRLDE 600	
QY	601 LREAMPNVNIQMLLRNRTVGYTPYDPSVCRAFVKEAASGVDFIRIFDALNDVQMRPA 660	
Db	601 LREAMPNVNIQMLLRNRTVGYTPYDPSVCRAFVKEAASGVDFIRIFDALNDVQMRPA 660	
QY	661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDMAGLLR 720	

Db	661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDMAGLLR 720	
QY	721 PAAVTKLVTLRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTSQPS 780	
Db	721 PAAVTKLVTLRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTSQPS 780	
QY	781 LSAIVAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840	
Db	781 LSAIVAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840	
QY	841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDALHLVGAGVDPADF 900	
Db	841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDALHLVGAGVDPADF 900	
QY	901 AADPOKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLLDA 960	
Db	901 AADPOKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLLDA 960	
QY	961 DDSKERNSLNRLLPKPTBFLHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020	
Db	961 DDSKERNSLNRLLPKPTBFLHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTP 1020	
QY	1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPEFA 1080	
Db	1021 LLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAPEFA 1080	
QY	1081 GVVTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRWVWPAATKVEGGDLIVVVS 1140	
Db	1081 GVVTVTVAEGDEVKAGDAVAIEAMKMEATITASVDGKIDRWVWPAATKVEGGDLIVVVS 1140	
RESULT 2		
Q8L2G4		
ID	Q8L2G4 PRELIMINARY; PRT; 1140 AA.	
AC	Q8L2G4;	
DT	01-OCT-2002 (TrEMBLrel. 22, Created)	
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Pyruvate carboxylase.	
GN	PYC.	
OS	Corynebacterium crenatum.	
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.	
OX	NCBI_TaxID=168810;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CD945;	
RA	Wang J., Ding J., Liu Y.;	
RT	"Cloning and Expression of Pyruvate Carboxylase Gene in	
RT	Corynebacterium crenatum CD945.;"	
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF503915; AAM27458.1; -	
DR	GO; GO:0005737; C:cytoplasm; IEA.	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0009374; F:biotin binding; IEA.	
DR	GO; GO:0016874; F:ligase activity; IEA.	
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.	
DR	GO; GO:0006094; P:gluconeogenesis; IEA.	
DR	GO; GO:0008152; P:metabolism; IEA.	
DR	InterPro; IPR001882; Biotin_BS.	
DR	InterPro; IPR005482; Biotin_carb_C.	
DR	InterPro; IPR000089; Biotin_lipoYL.	
DR	InterPro; IPR005479; Cbase_L_D2.	
DR	InterPro; IPR005481; Cbase_L_N.	
DR	InterPro; IPR000891; HMGL-like.	
DR	InterPro; IPR003379; PYC_OADA.	
DR	InterPro; IPR005930; Pyruv_carbox.	
DR	Pfam; PF02785; Biotin_carb_C; 1.	
DR	Pfam; PF00364; biotin_lipoYL; 1.	
DR	Pfam; PF00289; CPSase_L_chain; 1.	
DR	Pfam; PF02786; CPSase_L_D2; 1.	
DR	Pfam; PF00682; HMGL-like; 1.	

DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMs; TIGR01235; pyruv carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPSASE 2; 1.
SQ SEQUENCE 1140 AA; 123126 MW; FFA90BB7644C910E CRC64;

Query Match 99.5%; Score 5759; DB 2; Length 1140;
Best Local Similarity 99.3%; Pred. No. 5.1e-298;
Matches 1132; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVR 60
Db 1 MSTHTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVR 60

QY 61 IGTEGSPVKAYLDIDEIIGAARKVKADALYPGYGFLSENAQLARECAENGITFIGPTPEV 120
Db 61 IGTEGSPVKAYLDIDEIIGAARKVKADALYPGYGFLSENAQLARECAENGITFIGPTPEV 120

QY 121 LDLTGDKSRVTAARKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180
Db 121 LDLTGDKSRVTAARKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRF 180

QY 181 VASPDRLKLAETASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240
Db 181 VSSPDRLKLAETASREAAAFGDGAVYVERAVINPQHIEVQILGDRTGEVVHLYERDCS 240

QY 241 LQRRHQKVEIAPAQHLDPQLDRDICADAVKFCRSIGYQGAGTYVEFLVDEKGNHVFIEVN 300
Db 241 LQRRHQKVEIAPAQHLDPQLDRDICADAVKFCRSIGYQGAGTYVEFLVDEKGNHVFIEVN 300

QY 301 PRIQVEHTVTEEVTEVDLVKAQMRLLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360
Db 301 PRIQVEHTVTEEVTEVDLVKAQMRLLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNG 360

QY 361 FRPDTGTTITAYRSPGGAGVRLDGAALQGGBITAHFDSMLVKMTCRGSDFETAVARAQRAL 420
Db 361 FRPDTGTTITAYRSPGGAGVRLDGAALQGGBITAHFDSMLVKMTCRGSDFETAVARAQRAL 420

QY 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480
Db 421 AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDV 480

QY 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLRGSRDRLKQLGPAFAAFARDLREQDALAVTDTT 540
Db 481 TVNKPVGVRPKDVAAPIDKLPNIKDLPLRGSRDRLKQLGPAFAAFARDLREQDALAVTDTT 540

QY 541 FRDAHQSLLATRVRSFALKPAEAEVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600
Db 541 FRDAHQSLLATRVRSFALKPAEAEVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDE 600

QY 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFVKEAAASSGVDIIFRIFDALNDVSQMRPA 660
Db 601 LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFVKEAAATSGVDIIFRIFDALNDVSQMRPA 660

QY 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAREIVKSGAHILAIKDMAGLLR 720
Db 661 IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYYLKMAREIVKSGAHILAIKDMAGLLR 720

QY 721 PAAVTKLVTLALRREEDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780
Db 721 PAAATKLVTLALRREEDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS 780

QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840
Db 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQL 840

QY 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900
Db 841 SNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADF 900

QY 901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDA 960
Db 901 AADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDA 960

QY 961 DDSKERRNSLNRLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLRLPDVRTP 1020
Db 961 DDSKERRNSLNRLLPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLRLPDVRTP 1020

QY 1021 LLVRLDAISEPDDKGMNRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080
Db 1021 LLVRLDAISEPDDKGMNRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAPFA 1080

QY 1081 GWTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVVS 1140
Db 1081 GWTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAAATKVEGGDLIVVIS 1140

RESULT 3
Q8FRQ0
ID Q8FRQ0 PRELIMINARY; PRT; 1168 AA.
AC Q8FRQ0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1).
GN PYC OR CE0709.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005216; BAC17519.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001882; Biotin BS.
DR InterPro; IPR005482; Biotin carb C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PYC OADA.
DR InterPro; IPR005930; Pyruv_carbox.
DR Pfam; PF02785; Biotin carb C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMs; TIGR01235; pyruv carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPSASE 2; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 1168 AA; 126245 MW; A5DSA4DD4DF285F8 CRC64;

Query Match 91.9%; Score 5317.5; DB 16; Length 1168;
Best Local Similarity 91.2%; Pred. No. 1.7e-274;
Matches 1038; Conservative 43; Mismatches 56; Indels 1; Gaps 1;

QY 3 THTSSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIG 62
Db 32 TTTPTSTLPAPFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIG 91
QY 63 TEGSPVKAYLDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLD 122

Db	92	TEGSPVKAYLDIDEIINAAKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEVLD	151
QY	123	LTGDKSRAVTAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRFA	182
Db	152	LTGDKSAVSAAKKAGLPVLAESTPSTDIDEIVKSAEGQTYPIFVKAVAGGGRGMRFE	211
QY	183	SPDELKRLATEASREAEAAFGGAVYVERAVINPQHIEVQILGDHTGEVVLHYERDCSLQ	242
Db	212	KPEDLRELAREASREAEAAFGGGSVYVERAVIKPQHIEVQILGDHTGDVHLHYERDCSLQ	271
QY	243	RRHQKVVEIAPAQHLDPPELRDRI CADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPR	302
Db	272	RRHQKVVEIAPAQHLDPPELRDRI CADAVKFCRSIGYQGAGTVEFLVDEAGNHVFIEMNPR	331
QY	303	IQVEHTVTEEVTEVDLVKAQMRLAAGATLTKELGLTQDKIKTHGAALQCRITTEDPNNGFR	362
Db	332	IQVEHTVTEEVTSVDLVKAQMHLAAGATLTKELGLTQDKITTHGAALQCRITTEDPSNNFR	391
QY	363	PDTGITAYRSPGGAGVRLDGAAGLGGGEITAHFDSMLVKMTCRGSDFFETAVARAQRALAE	422
Db	392	PDTGVITAYRSPGGAGVRLDGAAGLGGGEITAHFDSMLVKMTCRGSDFFETAVSRAQRALAE	451
QY	423	FTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLQAPPADDEQGRILDYLDVTV	482
Db	452	FNVSGVATNIGFLRALLREEDFTKRRIDTGFIGSHQLLQAPPADDEQGRILEYLDVTV	511
QY	483	NKPHGVPRKDVAAPIDKLPNIKDPLPRGSRDRLKQLGPAAFARDLREQDALAVDTTFR	542
Db	512	NKPHGERP-ETARPIEKLPEVENIPLPRGSRDRLKQLGPEGFARDLREQDALAVDTTFR	570
QY	543	DAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRDLDEL	602
Db	571	DAHQSLLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDEL	630
QY	603	EAMPNVNIQMLLRNTVGYTPYDPSVCRAFYKAEAASSGVDIFRIFDALNDVQMRPAID	662
Db	631	EAMPNVNIQMLLRNTVGYTPYDPSVCRAFYQEAASKGVDIFRIFDALNDISQMRPAID	690
QY	663	AVLENTTAVAEVAMAYSGDLSDPNEKLYTLDYYLKNABEIVKSGAHILAIKDMAGLLRPA	722
Db	691	AVLETGTSVAEVAMAYSGDLSNPGEKLYTLDYYLNLAEQIVDSGAHILAIKDMAGLLRPA	750
QY	723	AVTKLVTALRREFDLPVHVHTHDTAGGQLATYFAAAQACADAVD GASAPLSGTSQPSLS	782
Db	751	AATKLVTALRREFDLPVHVHTHDTAGGQLATYLAANAGADAVDAASAPLSGTSQPSMS	810
QY	783	AIVAAFAHTRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN	842
Db	811	ALVAAFAHTRDRTGLNLQAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN	870
QY	843	LRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVDPADFAA	902
Db	871	LRAQAVALGLADRFELIEDYAAVNEMLGRPTKVTPSSKVVGDLALHLVGAGVSPEDFAA	930
QY	903	DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPPEEQAHLDADD	962
Db	931	DPQKYDIPDSVIAFLRGELGTPPGWPEPLRTRALEGRSQGKAPLAEIPAEQAHLDSD	990
QY	963	SKERNNSINRLLFPKPTEEFLEHRRRFGNTSALDDREFPYGLVEGRETLIRLPDVRTPLL	1022
Db	991	SAERRGTLNRLLFPKPTEEFLEHRRRFGNTSALDDREFPYGLKEGREELIRLTGVSTPMV	1050
QY	1023	VRLDAISEPDDKGMNVNVANVNGQIRPMRVDRDRSVESVTATAEKADSSNKGHVAAPFAGV	1082
Db	1051	VRLDAISEPDDKGMNVNVNVNNGQIRPIKVRDRSVESVTATAEKADATNKGHVAAPFAGV	1110
QY	1083	VTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS	1140
Db	1111	VTVTVAEGDEIKAGDAVAIIEMKMEATITAPVDGVVIDRVVVVPAATKVEGGDLIVVVS	1168

RESULT 4
Q8RQL2

ID	Q8RQL2	PRELIMINARY;	PRT;	1139	AA.
AC	Q8RQL2;				
DT	01-JUN-2002	(TReMBLrel. 21, Created)			
DT	01-JUN-2002	(TReMBLrel. 21, Last sequence update)			
DT	01-OCT-2003	(TReMBLrel. 25, Last annotation update)			
DE		Pyruvate carboxylase.			
GN	PYC.				
OS	Corynebacterium efficiens.				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.				
OX	NCBI_TaxID=152794;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Akiyoshi N., Nonaka G., Kimura E., Kawahara Y., Sugimoto S.;				
RT	"Corynebacterium efficiens pyruvate carboxylase (pyc) gene, complete				
RT	CDS."				
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.				
CC	-1- COFACTOR: BIOTIN (BY SIMILARITY).				
DR	EMBL; AB083299; BAB88903.1; -				
DR	GO; GO:0005737; C:cytoplasm; IEA.				
DR	GO; GO:0005524; F:ATP binding; IEA.				
DR	GO; GO:0009374; F:biotin binding; IEA.				
DR	GO; GO:0016874; F:ligase activity; IEA.				
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.				
DR	GO; GO:0006094; P:gluconeogenesis; IEA.				
DR	GO; GO:0008152; P:metabolism; IEA.				
DR	InterPro; IPR001882; Biotin_BS.				
DR	InterPro; IPR005482; Biotin_carb_C.				
DR	InterPro; IPR000089; Biotin_lipoY1.				
DR	InterPro; IPR005479; CPase_L_D2.				
DR	InterPro; IPR005481; CPase_L_N.				
DR	InterPro; IPR000891; HMGL-Like.				
DR	InterPro; IPR003379; PYC_OADA.				
DR	InterPro; IPR005930; Pyruv_carbox.				
DR	Pfam; PF02785; Biotin_carb_C; 1.				
DR	Pfam; PF00364; biotin_lipoY1; 1.				
DR	Pfam; PF00289; CPSase_L_chain; 1.				
DR	Pfam; PF02786; CPSase_L_D2; 1.				
DR	Pfam; PF00682; HMGL-Like; 1.				
DR	Pfam; PF02436; PYC_OADA; 1.				
DR	TIGRFAMS; TIGR01235; pyruv carbox; 1.				
DR	PROSITE; PS00188; BIOTIN; 1.				
DR	PROSITE; PS00867; CPSASE_2; 1.				
KW	Biotin.				
SQ	SEQUENCE	1139	AA;	123068	MW; BA7023134519FAAA CRC64;
Query Match					
Best Local Similarity		91.7%;	Score	5305.5;	DB 2; Length 1139;
Matches		1037;	Conservative	43;	Mismatches 59; Indels 1; Gaps 1;
QY	1	MSTHTSSTLP	PAFKKILVANRGEIAVRAFAALETGATGATVAIYPREDRGSFHRSFASEAVR	60	
Db	1	MVTTTPSTLP	PAFKKILVANRGEIAVRAFAAYETGATVAIYPREDRGSFHRSFASEAVR	60	
QY	61	IGTEGSPVKAYLDIDEIIGA	AKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEV	120	
Db	61	IGTEGSPVKAYLDIDEIINA	AKVKADAVYPGYGFLSENAQLARECAENGITFIGPTPEV	120	
QY	121	LDLTGDKSRAVTA	AAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGRGMRF	180	
Db	121	LDLTGDKSKAVSA	AKKAGLPVLAESTPSTDIDEIVKSAEGQTYPIFVKAVAGGGRGMRF	180	
QY	181	VASDELKRLATEAS	REAEAAFGDGAVVVERAVINPQHIEVQILGDHTGEVVLHYERDCS	240	
Db	181	VEKEDLRELAREAS	REAEAAFGDGSVVVERAVIKPQHIEVQILGDHTGDVHLHYERDCS	240	
QY	241	LQRHQKVVEIAPA	QHLDPELRDRI CADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMN	300	
Db	241	LQRHQKVVEIAPA	QHLDPELRDRI CADAVKFCRSIGYQGAGTVEFLVDEAGNHVFIEMN	300	
QY	301	PRIQVEHTVTEE	VTVDLVKAQMRLAAGATLTKELGLTQDKIKTHGAALQCRITTEDPNNG	360	
Db	301	PRIQVEHTVTEE	VTVDLVKAQMHLAGATLTKELGLTQDKITTHGAALQCRITTEDPSNN	360	

QY	361	FRPDTGTTITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVARAQAL	420
Db	361	FRPDTGVITAYRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRGSDFFETAVSRAQAL	420
QY	421	AEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADV	480
Db	421	AEFNVSGVATNIGFLRALLREEDFTKRRIDTGFIGHQHLLQAPPADDEQGRILEYADV	480
QY	481	TVNKPVGVRPKDVAAPIDKLPNIKOLPLPRGSRDRLKQLGPAAAFARDLREQDALAVTDIT	540
Db	481	TVNKPVGGERP-ETARPIEKLPEVENIPLPRGSRDRLKQLGPEGFARDLREQDALAVTDIT	539
QY	541	FRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLE	600
Db	540	FRDAHQSLLATRVRSFALTPAARAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWARLDE	599
QY	601	LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFVKEAASSGVDFRIFDALNDVSQMRPA	660
Db	600	LREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFVQEAAKSGVDIFRIFDALNDISQMRPA	659
QY	661	IDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAikomAGLLR	720
Db	660	IDAVLETGTSVAEVAMAYSGDLSNPGEKLYTLDYLLNLAEQIVDSGAHILAikomAGLLR	719
QY	721	PAAVTKLVTLRRFEDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPS	780
Db	720	RAAAPKLVTLRRFEDLPVHVHTHDTAGGQLATYLAANAGADAVDAASAPLSGTTSQPS	779
QY	781	LSAIVAAFAHTRRTDGLSLEAVSDLEPYWEAVRGLYLPFESGTPPTGRVYRHEIPGGQL	840
Db	780	MSALVAAFAHTRRTDGLNLQAVSDLEPYWEAVRGLYLPFESGTPPTGRVYRHEIPGGQL	839
QY	841	SNLRAQATALGLADRFELIEDNYAAVNEMLCRPTKVTTPSSKVVGDALHLVAGVDPADF	900
Db	840	SNLRAQAVALGLADRFELIEDYAAVNEMLCRPTKVTTPSSKVVGDALHLVAGVSPEDF	899
QY	901	AADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEQAHLDA	960
Db	900	AADPKYDIPDSVIAFLRGELGTPPGWPEPLRTRALEGRSQKAPLAEIPAEEQAHLDS	959
QY	961	DDSKERRNSLNRLLPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVSTP	1020
Db	960	DDSAERRGTLNRLLPKPTEEFLEHRRRFGNTSALDDREFFYGLKEGREELIRLTGVSTP	1019
QY	1021	LLVRLDAISEPDDKGMNRNVVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAPEA	1080
Db	1020	MVVRLLDAVSEPDKGMNRNVVNNGQIRPIKVRDRDSVESVTATAEKADATNKGHVAAPEA	1079
QY	1081	GVVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS	1140
Db	1080	GVVTVTVAEGDEIKAGDAVAIIEMKMEATITAPVDGVIDRVVVVPAATKVEGGDLIVVVS	1139
RESULT 5			
Q9RK64			
ID	Q9RK64	PRELIMINARY;	PRT; 1124 AA.
AC	Q9RK64;		
DT	01-MAY-2000	(TReMBLrel. 13, Created)	
DT	01-MAY-2000	(TReMBLrel. 13, Last sequence update)	
DT	01-OCT-2003	(TReMBLrel. 25, Last annotation update)	
DE		Pyruvate carboxylase.	
GN	SCO0546	OR SCF11.26C.	
OS		Streptomyces coelicolor.	
OC		Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC		Streptomycineae; Streptomycetaceae; Streptomyces.	
OX	NCBI_TaxID=1902;		
RN	[1]		
RP		SEQUENCE FROM N.A.	
RC		STRAIN=A3(2);	
RA		Seeger K.J., Harris D.;	
RL		Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.	
RN	[2]		
RP		SEQUENCE FROM N.A.	
RC		STRAIN=A3(2);	
RA		Barrell B.G., Rajandream M.A.;	
RL		Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.	
RN	[3]		
RP		SEQUENCE FROM N.A.	
RC		STRAIN=A3(2);	
RX		MEDLINE=97000351; PubMed=8843436;	
RA		Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,	
RA		Kinashi H., Hopwood D.A.;	
RT		"A set of ordered cosmids and a detailed genetic and physical map for	
RT		the 8 Mb Streptomyces coelicolor A3(2) chromosome.";	
RL		Mol. Microbiol. 21:77-96(1996).	
RN	[4]		
RP		SEQUENCE FROM N.A.	
RC		STRAIN=A3(2) / M145;	
RX		MEDLINE=21996410; PubMed=12000953;	
RA		Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,	
RA		Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,	
RA		Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,	
RA		Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,	
RA		Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,	
RA		Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,	
RA		Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,	
RA		Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,	
RA		Hopwood D.A.;	
RT		"Complete genome sequence of the model actinomycete Streptomyces	
RT		coelicolor A3(2).";	
RL		Nature 417:141-147(2002).	
DR		EMBL; AL939105; CAB59603.1; --	
DR		HSSP; P24182; 1BNC.	
DR		GO; GO:0005737; C:cytoplasm; IEA.	
DR		GO; GO:0005524; F:ATP binding; IEA.	
DR		GO; GO:0016874; F:ligase activity; IEA.	
DR		GO; GO:0004736; F:pyruvate carboxylase activity; IEA.	
DR		GO; GO:0006094; P:gluconeogenesis; IEA.	
DR		GO; GO:0008152; P:metabolism; IEA.	
DR		InterPro; IPR005482; Biotin_carb_C.	
DR		InterPro; IPR000089; Biotin_lipoyl.	
DR		InterPro; IPR005479; CPase_L_D2.	
DR		InterPro; IPR005481; CPase_L_N.	
DR		InterPro; IPR000891; HMGL-like.	
DR		InterPro; IPR003379; PYC_OADA.	
DR		InterPro; IPR005930; Pyruv_carbox.	
DR		Pfam; PF02785; Biotin_carb_C; 1.	
DR		Pfam; PF00364; biotin_lipoyl; 1.	
DR		Pfam; PF00289; CPSase_L_chain; 1.	
DR		Pfam; PF02786; CPSase_L_D2; 1.	
DR		Pfam; PF00682; HMGL-like; 1.	
DR		Pfam; PF02436; PYC_OADA; 1.	
DR		TIGRFams; TIGR01235; pyruv_carbox; 1.	
DR		PROSITE; PS00867; CPSASE_2; 1.	
KW		Complete proteome.	
SQ		SEQUENCE 1124 AA; 121071 MW; 1FEF4C4643954D31 CRC64;	
Query Match 65.6%; Score 3799; DB 16; Length 1124;			
Best Local Similarity 65.2%; Pred. No. 1.1e-193;			
Matches 736; Conservative 154; Mismatches 228; Indels 10; Gaps 6;			
QY	12	FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY	71
Db	2	FRKVLVANRGEIAIRAFAAGYELGARTVAVFPFHEDRNSLHRLKADEAYEIGEQQHPVRAY	61
QY	72	LDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEVLDLTGDKSRAY	131
Db	62	LSVEEIVRAARRAGADAVYPGYGFLSENPELARACEEAGITFVGPSARILELTGNKARAY	121
QY	132	TAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRVFVASPELRKLA	191
Db	122	AAAREAGVPVLGSSAPSTDVDDELVRADDVGFPPFVKAVAGGGGRGMRVVEEPAQLREAI	181
QY	192	TEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVHLYERDQSLQRRHQKVEI	251

Db 182 EAASREAAAFGSDTVFLEKAVVEPRHIEVQILADGEDGVIHLFERDCSVQRRHQKVIEL 241
QY 252 APAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEPNPRIQVEHTVTE 311
Db 242 APAPNLDPALRERICADAVNFARQIGYRNAGTVEFLVDRDGNHVFIEPNPRIQVEHTVTE 301
QY 312 EVTEVDLVKAQMLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY 371
Db 302 EVTDVDLVQSLRIAAGQTLDLGLAQENITLGAALQCRITTEDPANGFRPDTGTQISAY 361
QY 372 RSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRSGDFETAVARAQRALAEFTVSGVAT 430
Db 362 RSPGGGIRLDGGTTHAGTEISAHFDSMLVKLSRCGRDFTTAVNRRARRAVAEFRIRGVAT 421
QY 431 NIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDADVTVNKPHGVRP 490
Db 422 NIPFLQAVLDDPFOAGRVTTTSFIEQRPHLLTARHSADRGTKLLTYLADVTVNKPHGERP 481
QY 491 KDVAAPIDKLPNIXDLPPLGRSRDLKQLGPAFAFARDLREQDALAVTDTTFRDAHQSLLA 550
Db 482 -ELVDPLTKLPTASAGEPPAGSRQLLAEELGPGFARRLRESSTIGVTDTTFRDAHQSLLA 540
QY 551 TRVRSFALKPAEAAVAKLTPELLSVEAWGGATVDVAMRFLFEDPWDRDLDELREAMPVNI 610
Db 541 TRVRTKMLAVAPVAVARTLPQLLSLEWCWGATVDVALRFLAEDPWERLAALREAVENLCL 600
QY 611 QMLLRGRNTVGYTPYPSVCRAFVKEAASSGVDFIFRIDALNDVSMRPAIDAVLETNTA 670
Db 601 QMLLRGRNTVGYTPYPTVETDAFVQEAATGIDIFRIDALNDVEQMRPAIEAVRQTGSA 660
QY 671 VAEVAMAYSGDLSDPNEKLYTLDYYLKMAREEIVKSGAHILAIDKMGALLRPAAVTKLVTA 730
Db 661 VAEVALCYTADLSDPSERLYTLDYYLRLAEQIVNAGAHVLAVKDMAGLLRAPAAATLVSA 720
QY 731 LRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLSAIVAAFAH 790
Db 721 LRREFDLPVHLHTHTTTGGQLATYLAIIQAGADAVDGAVASMAGTTSQPSLSAIVAATDH 780
QY 791 TRRDTGLSLEAVSDLEPYWEAVRGLYLPFFESGTPPTGRVYRHEIPGGQLSNLRQAOTAL 850
Db 781 TERPTGLDLQAVGDLEPYWESVRKYAPFEAGLASPTGRVYHHEIPGGQLSNLRTOAVAL 840
QY 851 GLADRFELIEDNYAAVNMELGRPTKVTPSSKVVGDIALHLVGAGVDPADFAADPKYDIP 910
Db 841 GLGDRFEDIEAMYAADRMGLRLVKVTPSSKVVGDIALHLVGAGVSPADFEQDPDRFDIP 900
QY 911 DSVIAPLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPBEEQAHLDADDSSKERRNSL 970
Db 901 DSVVGLRGELGTPPGWPEBFRSKALRGAEAR-PLAELSEDDRDLG-----GKDRRATL 955
QY 971 NRLLFPKPTEEFLEHRRRRFNGTNSALDDREFFYGLVEGRETLIRL-PDVTRTPLLRLDAIS 1029
Db 956 NRLLFPGPAREFDTHRASVYGDTSILDSKOFFYGLRPGKEYTVDLDPGVR--LLIELQAVG 1013
QY 1030 EPDDKGMNRNVANVNGQIRPMRVDRSVESVTATAEKADSSNKGHVAAAPFAGVTVTVVAE 1089
Db 1014 DADERGMRTVMSSSLNGQLRPIQVRDRSAATDVPTTEKADRANPGHVAAAPFAGVTVLVAE 1073
QY 1090 GDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAATKVEGGDLIV 1137
Db 1074 GDEVEAGATVATIEAMKMEASITAPKSGTVRTLAINRIQQVEGGDLLV 1121
RESULT 6
Q9F843 ID Q9F843 PRELIMINARY; PRT; 1127 AA.
AC Q9F843;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1).
GN PYC.
OS Mycobacterium smegmatis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20374587; PubMed=10913817;
RA Mukhopadhyay B., Purwantinil E.;
RT "Pyruvate carboxylase from Mycobacterium smegmatis: stabilization,
RT rapid purification, molecular and biochemical characterization and
RT regulation of the cellular level";
RL Biochim. Biophys. Acta 1475:191-206(2000).
CC -I- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; AF262949; AAG30411.1; -.
DR HSSP; P24182; 1BNC.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PYC_OADA.
DR InterPro; IPR005930; Pyruv_carbox.
DR InterPro; IPR000634; S/T dehydratse_BS.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPASE_2; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Biotin; Ligase.
SQ SEQUENCE 1127 AA; 120765 MW; 18C132C48425C67B CRC64;
Query Match 65.4%; Score 3783.5; DB 2; Length 1127;
Best Local Similarity 66.2%; Pred. No. 7.7e-193;
Matches 747; Conservative 141; Mismatches 234; Indels 7; Gaps 6;
QY 14 KILVANRGEIAVRAFRALLETGAATVAIYPREDRSGFHRSFASEAVRIGTEGSPVKAYLD 73
Db 4 KVLVANRGEIAIRAFRAAYEMGIATVAVPYPERNSLRLKADESQIGEVGHPVAYLS 63
QY 74 IDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAVTA 133
Db 64 VDEIRVAKHSGADAVYPGYGFLSENPDLAACKCAEAGITFVGPSAEVLQLTGNKARAIAA 123
QY 134 AKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELRLKATE 193
Db 124 ARAAGLPVLSSESPSSVDELMAAAADMEFFLFVKAVSGGGGRMRRVTRDRESLAEIAEA 183
QY 194 ASREABAAFGDGAIVVERAVINPOHIEVQILGDHTGEVVLHYERDCSLQRRHQKVEIAP 253
Db 184 ASREAESAFGDASVYLEQAVLNPRHIEVQILADGAGNVMLFERDCSVQRRHQKVELAP 243
QY 254 AQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEPNPRIQVEHTVTEEV 313
Db 244 APNLSDELRRQQICADAVAFARQIGYSCAGTVEFLDGRGHVFIENPRIQVEHTVTEEI 303
QY 314 TEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAYRS 373
Db 304 TDVDLVSSQLRIAAGETLADLGLSQDRLLVVRGAAMQCRITTEVPANGFRPDTGTITAYRS 363
QY 374 PGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRSGDFETAVARAQRALAEFTVSGVATNIG 433

Db 364 PGGAGIRLDGNTLGAIEISAHFDSMLVKLTCRGRDFAAASRRARRALAEFRIRGVSTNIP 423
Qy 434 FLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYADVTVNKPHGVPRKDV 493
Db 424 FLOAVIDDPDFRAGRVTTSFIDDRPHLLTSRSPADRGTRILNYLADITVNKPHGERPSTV 483
Qy 494 AAPIDKLPNIKDL--PLPRGSRDLKQLGPAFAFARDLREQDALAVTDTTFRDAHQSLLAT 551
Db 484 -YPQDKLPL-LDQAPPPAGSKORLVELGPEGFAGWLRESKAVGVTDTTFRDAHQSLLAT 541
Qy 552 RVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPDWRDLDELREAMPNNIQ 611
Db 542 RVRTTGLLMVAPYVARSPQLLSIECWGATYDVALRFLKEDPWERLAALRESVPNICLQ 601
Qy 612 MLLGRNTVGYTPYDPSVCRAFKVEAASSGVDFRIFDALNDVSMRPAIDAVLETNTAV 671
Db 602 MLLGRNTVGYTPYDELVTSAFVEEAAATGIDIFRIFDALNNVESMRPAIDAVREITSTI 661
Qy 672 AEVAMAYSGDLSDPNEKLYTLDYLLKMAEIEIVKSGAHILAIKDMAGLLRPAAVTKLVATL 731
Db 662 AEVAMCYTGDLSPAENLYTLDYLLKLAQIVEAGAHVLAIKDMAGLLRPAHAHTLVSA 721
Qy 732 RREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDVGASAPLSGTSQPSLSAIVAAFAHT 791
Db 722 RSRFDLPVHVHTHTPGGQLATYLAAWSAGADAVDVGASAPMAGITSQPALSIVAAAAHT 781
Qy 792 RRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGQSLNLRQAATAG 851
Db 782 QYDTGLDLRAVCDLEPYWEAVRKVYAPFESGLPGPTGRVYTHEIPGQSLNLRQAIALG 841
Qy 852 LADRFEIENYAAVNEMLGRPTKVTPTSSKVVGDALHLVGAGVDPADFAADPKYDIPD 911
Db 842 LGDRFEEIEANYAAADRVLGRVKTPTSSKVVGDALALVGAGITAEFEPAEDPAKYDIPD 901
Qy 912 SVIAFLRGELGNPPGWPPEPLRTRALEGRSEKAPLTEVPEEEOAHLDDADDKERNLSN 971
Db 902 SVIGFLRGELGDPGWPPEPLRTKALQGRGPAP-PVEKLTADDEALL-AQPGPKRQAALN 959
Qy 972 RLLFPKPTTEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEP 1031
Db 960 RLLFPGPATAEFAHRETYGDTSSLSANQFFGLRYGEEHRVQL-ERGVELLIGLEAISEA 1018
Qy 1032 DDKGMNVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAPFAGVTVTTVAEGD 1091
Db 1019 DERGMRTVMCIINGQLRPVLRDRDSIASEVPAEAKADRNNADHIAAPFAGVTVTVGAEGD 1078
Qy 1092 EVKAGDAVAIEAMKMEATITASVDGKIDRVVVPAATKVEGGDLVVVS 1140
Db 1079 SVDAGQTIATIEAMKMEAAITAPKAGTVARVAATAQVEGGDLVVVS 1127

RESULT 7
P95127 PRELIMINARY; PRT; 1127 AA.
ID P95127
AC P95127;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein (Pyruvate carboxylase).
GN PCA OR RV2967C OR MTCY349.20 OR MT3045.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z83018; CAB05410.1; --
DR PIR; D70671; D70671.
DR HSSP; P24182; 1BNC.
DR TIGR; MT3045; --
DR Tuberculist; RV2967C; --
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoYL.
DR InterPro; IPR005479; C:ase_L_D2.
DR InterPro; IPR005481; C:ase_L_N.
DR InterPro; IPR000891; HMGL-Like.
DR InterPro; IPR003379; PYC_OADA.
DR InterPro; IPR005930; Pyruv carbox.
DR InterPro; IPR000634; S/T dehydrase_BS.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoYL; 1.
DR Pfam; PF00289; CPSase_L chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMS; TIGR01235; pyruv carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00867; CPSASE 2; 1.
DR PROSITE; PS00165; DEHYDRATASE SER_THR; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1127 AA; 120422 MW; 84B0A4CC1A23CD90 CRC64;
Query Match 64.1%; Score 3712.5; DB 16; Length 1127;
Best Local Similarity 64.4%; Pred. No. 4.6e-189;
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;
Qy 12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
Db 2 FSKVLVANRGEIAIRAFAAYELGVGTVAVYPYEDRNSQHLKADESQIGDHPVHAY 61
Qy 72 LDIDEIIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRV 131
Db 62 LSVDEIVATARRAGADAIYPGYGFLSENPDIAAACAAGISFVGPSAEVLELAGNKSRAI 121
Qy 132 TAAKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVAPDELRLKLA 191
Db 122 AAAREAGLPVLMSSAPSASVDELLSVAAGMPFPLFVKAVAGGGGRMRRVGDIAALPEAI 181
Qy 192 TEASREAEAAFGDGAIVYVERAVINPQHIEVQILGHTGEVWHLYERDCSLQRHOKVVEI 251
Db 182 EAASREAESAFGDPVTYLEQAVINPRHIEVQILADNLGDLVIHLIERDCSVQRHOKVIEL 241

QY	252	APAQHLDP	ELDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIE	MNPRIQVEHTVTE	311
Db	242	APAPHLDAELRYKMCVD	AVAFARHIGYSCAGTVEFLD	ERGEYVFIEMNPRVQVEHTVTE	301
QY	312	EVTEVDLVKAQMR	LAAGATLKELGLTQDKIKTHGAALQCRITTEDP	NNGFRPDTGTITAY	371
Db	302	EITDVLVASQLRIA	AGETLEQLGRQEDIA	PHGAALQCRITTEDPANGFRPDTGRISAL	361
QY	372	RSPGGAGVRLDGA	AQLGGEITAHEDSMLVKMTCRGSDFETAVARAQALAEFTV	SGVATN	431
Db	362	RTAGGAGVRLDGS	TNLGAEISPYFDSMLVKLT	CRGRDLPTAVSRARRAIAEFRI	RGVSTN 421
QY	432	IGFLRALLREED	FTSKRIATGFIADHPHLLQAPPADDEQGRILDY	LADVTVNKPHGVRPK	491
Db	422	IPFLQAVLDDPD	FRAGRVTTSFIDERPQLLTARASADRGT	KILNFLADVTVNNPYGSRPS	481
QY	492	DVAAPIDKLP	NIKOLPL----PRGSRDRLKQLGPAAFARDLREQD	ALAVTDTTFRDAHQS	547
Db	482	TI-YPDDKLP	--DLDLRAAPPAGSKQRLVKLGPEGFARWLRESA	AVGVTDTTFRDAHQS	537
QY	548	LLATRVRSFALK	PAEAAVAKLTPELLSVEAWGGATYDVAMRFLFED	PWDRLDLREAMPN	607
Db	538	LLATRVRTSGLS	RVAPYLARTMPQLLSVECWGGATYDVALRFLKED	PWERLATLRAAMPN	597
QY	608	VNIQMLLRGRNT	VGTYTPYSDSVCRAFVKEAASSGVDFIFRIFDAL	NDVSQMRPAIDAVLET	667
Db	598	ICLQMLLRGRNT	VGTYTPYPEIVTSFVQEAATATGIDIFRIFDAL	NNIESMRPAIDAVRET	657
QY	668	NTAVAEVAMAYS	GDLSDPNEKLYTDYLYLKMAEEIVKSGAHILAIK	DMAGLLRPAAVTKL	727
Db	658	GSAIAEVAMCYT	GDLTDPGEQLYTDYLYLKLAEQIVDAGAHVLAIK	DMAGLLRPPAAQRL	717
QY	728	VTALRREFDL	LPVHVHTHTAGGQLATYFAAAQAGADAVDGASAP	LSGTTSQPSLSAIVAA	787
Db	718	VSALRSRFDLP	VLHLTHDTPGQLASYVAAWHAGADAVDGAAP	LAGTTSQPALSSIVAA	777
QY	788	FAHTRRDTGLS	LEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYR	HEIPGQLSNLRAQA	847
Db	778	AAHTEYDTGLS	LSAVCALEPYWEALRKVYAPFESGLPGPTGRVY	HEIPGQLSNLRQA	837
QY	848	TALGLADRFE	LIEDNYAAVNEMLGRPTKVT	PSSKVYGDALHLVGAGVDPADFAADPQKY	907
Db	838	IALGUDRFE	EIEEAYAGADRVLGRLVKTPTSKVYGDALALV	GAGVSADFEASDPARF	897
QY	908	DIPDSVIAFL	RGELGNPPGWPPEPLRTRALEGRSEKAPL	TEVPEEEQAHLDADDSKERR	967
Db	898	GIPESVLGR	ELGELDPPGGWPEPLRTAALAGRGAAR-PTA	QLAADDEIALSSVGAK-RQ	955
QY	968	NSLNFLLFPK	PTEEFLEHRRRFGNTSALDDREFFYGLV	EGRETILRLPDVRTPLLVRLDA	1027
Db	956	ATLNFLLFP	SPPTKEFNEHREAYGDTSQLSANOFFYGLR	QGBEHRVKL-ERGV	ELLIGLEA 1014
QY	1028	ISEPDDKGM	NVANVNGQIRPMRVDRSVESVTATAEKADSS	NKGHVAAPFAGVVTIVT	1087
Db	1015	ISEPDERGM	RTVMCILNGQLRPVLVRDRSIA	SAVPAEAKADRGNP	GHIAAPFAGVVTIVG 1074
QY	1088	AEQDEVKAG	DVAIIEMKMEATITASVDGKIDRVVVP	PAATKVEGGDLIVVVS	1140
Db	1075	CVGERVAG	QGTIATIEAMKMEAPITAPVAGTVER	VAVSDTAQVEGGDLLVVVS	1127
RESULT 8					
Q7TXJ1					
ID	Q7TXJ1	PRELIMINARY;		PRT;	1127 AA.
AC	Q7TXJ1;				
DT	01-OCT-2003	(TrEMBLrel. 25, Created)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	Probable pyruvate carboxylase PCA (Pyruvic carboxylase)				
DE	(EC 6.4.1.1).				
GN	PCA OR MB2991C.				
OS	Mycobacterium bovis				

OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX	NCBI_TaxID=1765;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=AF2122/97;
RX	MEDLINE=22709107; PubMed=12788972;
RA	Garnier T., Eigelmeier K., Camus J.-C., Medina N., Mansoor H.,
RA	Pryor M., Duthoy S., Lacroix C., Monsempe C., Simon S.,
RA	Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA	Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT	"The complete genome sequence of Mycobacterium bovis.";
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR	EMBL; BX248344; CAD96678.1; --
KW	Ligase; Complete proteome.
SQ	SEQUENCE 1127 AA; 120423 MW; 84B0A4CCL1A23CD90 CRC64;
Query Match 64.1%; Score 3712.5; DB 16; Length 1127;	
Best Local Similarity 64.4%; Pred. No. 4.6e-189;	
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;	
QY	12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY 71
Db	2 FSKVLVANRGEIAIRAFAAYELGVTVAVIYPYEDRNSQHRLLKADESYQIGDIGHPVHAY 61
QY	72 LDIDEIIGAARKVADAIYPGYGFLSENAQLARECAENGITFIGTPEVLDLTGDKSRV 131
Db	62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSAEVLELAGNKSRAI 121
QY	132 TAACKAGLPVLAESTPFSKNIDEIVKSAEQTYPIFKAVAGGGGRGMREFVASPDELRLKLA 191
Db	122 AAAREAGLPVLMSSAPSASVDELLSVAAGMFPFLFKAVAGGGGRGMRRVGDIAALPEAI 181
QY	192 TEASREAEAAFGDGVYVERAVINPQHIEVQILGDHTGEVHLYERDCSLQRRHQKVEI 251
Db	182 EAASREAESAFGDPVTVYLEQAVINPRHIEVQILADNLGDVIHLYERDCSVQRRHQKVI 241
QY	252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIE
Db	242 APAPHLDAELRYKMCVD
QY	312 EVTEVDLVKAQMR
Db	302 EITDVLVASQLRIA
QY	372 RSPGGAGVRLDGA
Db	362 RTAGGAGVRLDGS
QY	432 IGFLRALLREED
Db	422 IPFLQAVLDDPD
QY	492 DVAAPIDKLP
Db	482 TI-YPDDKLP
QY	548 LLATRVRSFALK
Db	538 LLATRVRTSGLS
QY	608 VNIQMLLRGRNT
Db	598 ICLQMLLRGRNT
QY	668 NTAVAEVAMAY
Db	658 GSAIAEVAMCYT
QY	728 VTALRREFDL
Db	718 VSALRSRFDLP

QY 788 FAHTRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPQPTGRVYRHEIPGGQLSNLRAQA 847
Db 778 AAHTEYDTGLSLSAVCALEPYWEALRVYAPFESGLPGPTGRVYHHEIPGGQLSNLQQQA 837
QY 848 TALGLADRFELIEDNYAAVNMELGRPTKVTSSKVVGDALHLVAGVDPADPAADPKY 907
Db 838 IALGLGDRFEEIEEAYAGADRVGLRVKVTPTSKVVGDALALVAGVSADEFASDPARF 897
QY 908 DIPDSVIAFLRGELGNPPGGWPEPLRTALRSEGRSEKAPLTVPEEEQAHLDDADDSKERR 967
Db 898 GIPESVLGFLRGELGDPGGWPEPLRTAALAGRGAR-PTAQLAADDEIALSSVGAK-RQ 955
QY 968 NSLNRLFPKPTTEEFLEHRRRRCNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
Db 956 ATLNRLLFPSPPTKEFNEHREAYGDTSQLSANOFFYGLRQGEHRVKL-ERGVELLIGLEA 1014
QY 1028 ISEPDDKGMNVVANVNGOIRPMVRDRSVESVTATAEAKADSSNKGHVAAPFAGVVTTV 1087
Db 1015 ISEPDERGMRVTMCIILNGQLRFLVLRDRSIASAVPAEAKADRGNPGHIAAPFAGVVTGV 1074
QY 1088 AEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVVPAATKVEGGDLVVVS 1140
Db 1075 CVGERVAGQTIATIEAMKMEAPITAPVAGTVERVAVSDTAQVEGGDLLVVVS 1127

RESULT 9
Q50450
ID Q50450 PRELIMINARY; PRT; 1124 AA.
AC Q50450;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Pyc.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: BIOTIN (BY SIMILARITY).
DR EMBL; U00024; AAA50948.1; -.
DR HSSP; P24182; 1BNC.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0009374; F:biotin binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR000891; HMGL-Like.
DR InterPro; IPR003379; Pyc_OADA.
DR InterPro; IPR005930; Pyruv_carbox.
DR InterPro; IPR000634; S/T dehydrtse_BS.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF00682; HMGL-Like; 1.
DR Pfam; PF02436; Pyc_OADA; 1.
DR TIGRFAMs; TIGR01235; pyruv carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.

DR PROSITE; PS00867; CPSASE 2; 1.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Biotin.
SQ SEQUENCE 1124 AA; 120672 MW; 07E1692E12203491 CRC64;
Query Match 60.2%; Score 3483; DB 2; Length 1124;
Best Local Similarity 62.2%; Pred. No. 7.6e-177;
Matches 698; Conservative 151; Mismatches 251; Indels 22; Gaps 9;
QY 12 FKKILVANRGEIAVRAFRAALETGAATVAIYPREDRGSEFHRSEFASEAVRIGTEGSPVKAY 71
Db 2 FSKVLVANRGEIAIRAFAAYELGVGTVAVYPYEDRNSQHLKADESQIGDIGHPVHAY 61
QY 72 LDIDELIGAARKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRV 131
Db 62 LSVDEIVATARRAGADAIYPGYGFLSENPDLAAACAAAGISFVGPSEVLELAGNKSRAI 121
QY 132 TAAKKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDRLKLA 191
Db 122 AAAREAGLPVLMSSAPASVDLSSVAAGMPFPLFVKAVAGGGGRGMRRVGDIAALPEAI 181
QY 192 TEASREAAAFDGAIVYVERAVINPQHIEVQILGDHTCEVHLYERDCSLQRRHQKVEI 251
Db 182 EAASREAESAFGDPVTVYLEQAVINPRHIEVQILADNLGDVILHYERDCSVQRRHQKVIEL 241
QY 252 APAQHLDPELDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEMNPRIOVEHTVIE 311
Db 242 APAPHLDAELRYKMCVDVAFARHIGYSCAGTVEFLDERGEYVFIEMNPRVQVEHTVIE 301
QY 312 EVTEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFRPDGTGITAY 371
Db 302 EITDVDLVASQLRIAAGETLEQLGRQEDIAHPGAALQCRITTEDPANGFRP-TRAGSAR 360
QY 372 RSPGGAGVRLDGAALQGEITAHFDSMLVKM-----TCRSDFFETAVARAQALAEFTVS 426
Db 361 CDPPAVPVSAWTAAP-----TWRRNQPVRLRHAGQADLSGRDLPTAVSRARRAIAEFRIR 415
QY 427 GVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDVTVNKP 486
Db 416 GVSTNIPFLQAVLDDPDFRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPY 475
QY 487 GVRPKDVAAPIDKLPNIKOLPL-----PRGSRDLKQLGPAAPARDLREQDALAVTDTTFR 542
Db 476 GSRPSTI-YPDDKLP---DLDLRAAPPAGSKQRLVKLGPEGARLWLRSAAVGVTDTTFR 531
QY 543 DAHQSLLATRVRSFALKPAAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRLDEL 602
Db 532 DAHQSLLATRVRTSGLSRVAPYLARTMPQLLSVECWGGATYDVALRFLKEDPWERLATLR 591
QY 603 EAMPNVNIQMLLRGRNTVGYTPYDSVCRAFYKAEAASSGVVDIFRIFDALNDVSOQMRPAID 662
Db 592 AAMPNICLQMLLRGRNTVGYTPYPEIVTSFAVQEATATGIDIFRIFDALNNIESMRPAID 651
QY 663 AVLETNTAVAEVAMAYSGLSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIDMAGLLRPA 722
Db 652 AVRETGSAIAEVAMCYTGDLTDPGEQLYTLDYLYKLAEQIVDAGAHVLAIKDMAGLLRPP 711
QY 723 AVTKLVTLARRRFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASAPLSGTTSQPSLS 782
Db 712 AAQRLVSALRSRFDLPVHLHTHDTPGGQLASVVAWAHAGADAVDGAAPLAGTTSQPALS 771
QY 783 AIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSN 842
Db 772 SIVAAAHAHTEYDTGLSLSAVCALEPYWEALRVYAPFESGLPGPTGRVYHHEIPGGQLSN 831
QY 843 LRAQATALGLADRFELIEDNYAAVNMELGRPTKVTSSKVVGDALHLVAGVDPADFAA 902
Db 832 LRQQAIALGLGDRFEEIEEAYAGADRVGLRVKVTPTSKVVGDALALVAGVSADEFAS 891
QY 903 DPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTALRSEGRSEKAPLTVPEEEQAHLDDAD 962
Db 892 DPARFGIPESVLGRGELGDPGGWPEPLRTAALAGRGAR-PTAQLAADDEIALSSVG 950

Db	1065	IASPFAGQVTIKVDGVDEVSGQAVAILKMKYTTVNA	PVSGQVIRISIPPGRQVDIGD	1124
Qy	1135	LIVVV	1139	
Db	1125	LIMEI	1129	
RESULT 11				
Q83FS5	ID	Q83FS5	PRELIMINARY;	PRT; 1131 AA.
AC	Q83FS5;			
DT	01-JUN-2003	(TReMBLrel. 24, Created)		
DT	01-JUN-2003	(TReMBLrel. 24, Last sequence update)		
DT	01-OCT-2003	(TReMBLrel. 25, Last annotation update)		
DE	Pyruvate carboxylase (EC 6.4.1.1).			
GN	PCA OR TWT630.			
OS	Tropheryma whipplei (strain Twist) (Whipple's bacillus).			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Micrococcineae; Cellulomonadaceae; Tropheryma.			
OX	NCBI_TaxID=203267;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Raoult D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,			
RA	Claverie J.-M.;			
RT	"Tropheryma whipplei illustrates the diversity of gene loss patterns			
RT	in small genome bacterial pathogens.";			
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AE016852;	AAO44727.1;		
DR	GO; GO:0005737;	C:cytoplasm; IEA.		
DR	GO; GO:0004190;	F:aspartic-type endopeptidase activity; IEA.		
DR	GO; GO:0005524;	F:ATP binding; IEA.		
DR	GO; GO:0009374;	F:biotin binding; IEA.		
DR	GO; GO:0016874;	F:ligase activity; IEA.		
DR	GO; GO:0004736;	F:pyruvate carboxylase activity; IEA.		
DR	GO; GO:0006094;	P:gluconeogenesis; IEA.		
DR	GO; GO:0008152;	P:metabolism; IEA.		
DR	GO; GO:0006508;	P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR001969;	Aspprotease_AS.		
DR	InterPro; IPR001982;	Biotin_BS.		
DR	InterPro; IPR005482;	Biotin_carb_C.		
DR	InterPro; IPR000089;	Biotin_lipoY1.		
DR	InterPro; IPR005479;	CPase_L_D2.		
DR	InterPro; IPR005481;	CPase_L_N.		
DR	InterPro; IPR000891;	HMGL-like.		
DR	InterPro; IPR003379;	PYC_OADA.		
DR	InterPro; IPR005930;	Pyruv_carbox.		
DR	Pfam; PF02785;	Biotin_carb_C; 1.		
DR	Pfam; PF00364;	biotin_lipoY1; 1.		
DR	Pfam; PF00289;	CPSase_L_chain; 1.		
DR	Pfam; PF02786;	CPSase_L_D2; 1.		
DR	Pfam; PF00682;	HMGL-like; 1.		
DR	Pfam; PF02436;	PYC_OADA; 1.		
DR	TIGRFAMS; TIGR01235;	pyruv_carbox; 1.		
DR	PROSITE; PS00141;	ASP_PROTEASE; 1.		
DR	PROSITE; PS00188;	BIOTIN; 1.		
DR	PROSITE; PS00867;	CPSASE_2; 1.		
KW	Ligase; Complete proteome.			
SQ	SEQUENCE 1131 AA; 124300 MW; 6C21F973DA524CB1 CRC64;			
Query Match				
Best Local Similarity 56.0%; Score 3241; DB 16; Length 1131;				
Matches 668; Conservative 152; Mismatches 291; Indels 34; Gaps 10;				
Qy	12	FKXILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAY	71	
Db	2	FKXILIANRGEIAIRISRAAPERQIQTVAIYAHEDRNSLHRLKADEAYQIQGVGSPVAAY	61	
Qy	72	LDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKSRAY	131	
Db	62	LDVHEIRVALMSNSDAIHPGYGFLSEYLLADEAEKNGITFIGPPKDVLSAGDKVLAK	121	
Qy	132	TAAKAGLPVLAESTPSKNIDEIVKSAEQTYPIFVKAVAGGGGRGMRFVASPDRLKLA	191	

RESULT 12
Q8JHF6

Db	122	HMAQAAGLPTLRSSTASSNYDELLREAELEPYPIFVKAASGGGGRMIRVENRTALKNSL	181
Qy	192	TEASREAEAAFGDGAVVYVERAVINPOHIEVOILGDHTGEVVHLYERDCSLQRRHQKVEI	251
Db	182	ESAIQEAASAFGDRVFLETALDKPRHIEVQVLADKFGNIVHLFERDCSLQRRHQKVEI	241
Qy	252	APAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEHNPRIQVEHTVTE	311
Db	242	APAPNIPELLRTTLVRDAIAFAKSVKYENAGTVEFLVDSKMHYFIEHNPRIQVEHTVTE	301
Qy	312	EVTEVDLVKAQMRLLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITAY	371
Db	302	EITDIDIVOSQIILAAASLDDIGLVQDKIERRGFALQCRITTEDPHANFRPDTGRITSY	361
Qy	372	RSPGGAGVRLDGAAR-QLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTVSGVAT	430
Db	362	QSPGGAGIRLDASAVNPGVEITPYFDSMLVKMTCRGNFADATNRARRGLAEFRVGGVAT	421
Qy	431	NIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEOQRIILDYLDVTVNKPVGVRP	490
Db	422	NISFLRTLDDSSNFLNADFTTFIEKNGYLLQQSFLLDKHDRLSVYLGHVTVNKPYGDRP	481
Qy	491	KDVAAPIDKLPNIKDLPLP-----RGRDRLKQLGPAAFARDLREQDALAVTDTTF	541
Db	482	-ELVDPCSKIANF----LPDSSEVKKSVEGRDVLRLGPGQFAKLLCRKGLAVTDTTF	536
Qy	542	RNAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRDEL	601
Db	537	RNAHQSLLATRVRTIDLSRAAECTSAALPELSMEVWGATYDVALRELYEDPWERLSKI	596
Qy	602	REAMPNVNIQMLLRGRNTVGYTPYPDSVCRAFVKEAASSGVDIFRIFDALNDVSMRPAI	661
Db	597	REKVSICLQMLLRGRNTVGYTPYDQVTRAFVDEASDLGIDIFRIFDALNDVDMRIAI	656
Qy	662	DAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKMAEIVKSGAHILAKDMAGLLRP	721
Db	657	DAVQQTN-SVAEVAICYTGDLDDKRETVYTYIDYLEIAKKIVDAGAHILAKDMAGVLRP	715
Qy	722	AAVTKLVTLRRREFDLPVHVHTHTDTAGGQLATVFAAAQAGADAVDGASAPLSGTTSPSL	781
Db	716	RAATLLVSALKREFALPVHLHTDTPGGQLATLLAAADSGVDAVDVASGPMSGTTSPSM	775
Qy	782	SAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVVRHIEIPGGQLS	841
Db	776	SSLVAATDNTETHTGLSLSRVNELEPYWEAVRRLYVPFESGLLSPTGRVYIHEIPGGQLS	835
Qy	842	NLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVDPADFA	901
Db	836	NLKQQAIALGLSDRFEIIEEMYAYVNTLFGRIPKVTTPSSKVVGDLALYLASVNPDLGDFE	895
Qy	902	ADPKYDIPDSVIAFLRGELGNPPGGWPEPLRTPALEGRSEKAPLTEVPEEEQAHLDDAD	961
Db	896	MNPKKYDIPDSVISFLAGELGTPPAGWPD-FRDRVLAER-----EI-SIEQHPLSSD	945
Qy	962	DSKE-----RRNSLNRLLPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLLRL	1014
Db	946	DSKNLATSGKIRQQTLSKLLFPEPYRAFEANRAEYGDLSILRSEEFFYGLDFGIEYKIAV	1005
Qy	1015	PVVRTPLLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNKGH	1074
Db	1006	SS-SVGILVRLEAIGGVDSKGRSLVLSVNGELRPIQVRDESANVEVSRAEKADPNPGH	1064
Qy	1075	VAAPFAGVTVTVTAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAATKVEGGD	1134
Db	1065	IASPFAGQVTIKVDGVDEVVSGQAVAILKMKYTTVNA	1124
Qy	1135	LIVVV	1139
Db	1125	LIMEI	1129

ID	Q8JHF6	PRELIMINARY;	PRT;	1178 AA.
AC	Q8JHF6;			
DT	01-OCT-2002	(TrEMBLrel. 22, Created)		
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	Pyruvate carboxylase	(EC 6.4.1.1).		
GN	PYC.			
OS	Gallus gallus	(Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22145742; PubMed=12150961;			
RA	Jitrapakdee S., Nezis M.G., Cassidy A.I., Khew-Goodall Y.,			
RA	Wallace J.C.;			
RT	"Molecular cloning and domain structure of chicken pyruvate			
RT	carboxylase.";			
RL	Biochem. Biophys. Res. Commun. 295:387-393 (2002).			
DR	EMBL; AF509529; AAM92771.1; -			
DR	GO; GO:0005737; C:cytoplasm; IEA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0016874; F:ligase activity; IEA.			
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.			
DR	GO; GO:0006094; P:gluconeogenesis; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPR005482; Biotin carb C.			
DR	InterPro; IPR000089; Biotin_lipoyl.			
DR	InterPro; IPR005479; CPase_L_D2.			
DR	InterPro; IPR005481; CPase_L_N.			
DR	InterPro; IPR000891; HMGL-like.			
DR	InterPro; IPR003379; PYC_OADA.			
DR	InterPro; IPR005930; Pyruv carbox.			
DR	Pfam; PF02785; Biotin carb_C; 1.			
DR	Pfam; PF00364; biotin_lipoyl; 1.			
DR	Pfam; PF00289; CPase_L_chain; 1.			
DR	Pfam; PF02786; CPase_L_D2; 1.			
DR	Pfam; PF00682; HMGL-like; 1.			
DR	Pfam; PF02436; PYC_OADA; 1.			
DR	TIGRFAMs; TIGR01235; pyruv_carbox; 1.			
KW	Ligase.			
SQ	SEQUENCE	1178 AA; 127253 MW; F0722FEAB8BF39A5 CRC64;		
Query Match 44.1%; Score 2552; DB 13; Length 1178;				
Best Local Similarity 48.1%; Pred. No. 3.1e-127;				
Matches 556; Conservative 165; Mismatches 410; Indels 26; Gaps 14				
QY	2	STHTSSTLPAPKKILVANRGEIAVRAAFRAALETGAATVAIYPREDRGSFHRSPASEAVRI	61	
DB	28	SVRSASQP-IRKVLVANRGEIATRVFRACTELGLRTVAVYSEODTGMHRQKADAYLV	86	
QY	62	GTEGSPVKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEVL	121	
DB	87	GRGLPPVQAYLHVVDIIRVARENAVDAIHPGYGFLSERADFAQACVDAGRVFVGPPPEV	146	
QY	122	DLTGDKSRVTAAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMR	180	
DB	147	RKMGDKVEARSIAIAGVPVPGTSPVATLGEAQDFAARVGFPIIFKAAHGGGGGRGMR	206	
QY	181	VASPDRLKRLATEASREAEAAFGDGAUVVERAVINPQHIEVOILGDHTGEVVLVERDCS	240	
DB	207	VRGPQEEESFSRASSEALAAFGDGLFVEKLMERPRHIEGQILGDHGNVVLVERDCS	266	
QY	241	LQRRHQVVEIAPAQHLDPFLDRDICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMN	300	
DB	267	IQRHQVVEIAPAARLDPQLRAQLASDAVRIAQQVGYENAGTVEFLVDRDGKHYFIEVN	326	
QY	301	PRIQVEHTVTEEVTEVDLVKAQMRLLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG	360	
DB	327	SRLQVEHTVTEEITGVDLVQAQLLVAAGRSLSELGLQDQSVRVNGCAIQCRVTTEDPARG	386	
QY	361	FRPDTCTITAYRSPGGAGVRLDGAQL-GGEITAHFDSMLVKMTCRGSDSFETAVARAQRA	419	

Db		387	FQPDTRIEVFRSGEGMGIFLDCASAFQGALISPHYDSLLVKVIAHGPDQPSSAAAKMSRA	446
QY		420	LAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAAPPADDEQGRILDYLAD	479
Db		447	LGEFIRGVKTNIPFLQNVLAHFPQLGGAVDTQFIDENPELFHLRPPSONRAQKLLHYLG	506
QY		480	VTVNPKPHGVRP-KDVAAPIDKLPNIKDLPLPRGS-----RDRLKQLGPAAFAFDLRQDA	533
Db		507	VMVNGSPSTLPVKAKAAVVEPVPP---PVPMSGPPEGLRAVLQREGPAGFARALRGHRG	562
QY		534	LAVTDTTFRDAHQSLLATRVRSPALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFED	593
Db		563	LLLXDTTFRDAHQSLLATRVRTRDLARIAPEFAHSLSPCLCSMETWGGATFDVAMRFLHEC	622
QY		594	PWRDLBELREAMFNVINIOMLLGRNTVGTYTPYDPSVCRAFVKBAASSGVDIFRIFDALND	653
Db		623	PWERLRELRLVPNIPIPFQMLLRGANAVGYTNPDNVIYRFCEVAAAANGMDIFRIFDALNY	682
QY		654	VSQMRPAIDAVLENTATAEVAEMAYSGDSLSDPNEKLYTLDDYLLKMAEEIVKSGAHLAIK	713
Db		683	LPNLLLVGEAVGRAG-AVVEAALSVTGDVADPTRTKYSLDYVLGLAKELVAAGTHILCIK	741
QY		714	DMAGLLRPAAVTKLVIALRREF-DLPVHVHTHTTAGGQLATYFAAAQAGADAVDGASAPL	772
Db		742	DMAGLLTPAAARLLVSSLRRDRFPDPVPIHVHTHTAGAAIATLLAAANADADVWDVAVDAM	801
QY		773	SGTTSQPSLSAIVAAFAHTRRTDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPTGR--V	830
Db		802	SGMTSQPSMGALVACARGTPLDTGIALERVFHEYSEYEWEGARGLYAADFCTATMKSGNADV	861
QY		831	YRHEIPGGQSLNLRAQATALGLADRELIEDNYAAVNEMLGRPDKVTPSSKVVGDLALHL	890
Db		862	YENEIPGGQYNLHFQAHAMGLGHKFKEVKKAYAEANKLLGDLIKVTTPSSKVVGDLAQFM	921
QY		891	VGAGVDPADFAADPKYDIPDSVIAFLRGELGNPPGWPEPLRTRALEG--RSEGKAPLT	948
Db		922	VQNGLSREEABARADELSPFLSVVFELOQYIGTPPGGFPEPFRSKVLKDLPRVEGRPGAS	981
QY		949	EVPEEEQA---HLDADD--KERRNSLNRLFFPKPTEEBFLEHRRRRFGNTSALDDREFFYG	1003
Db		982	LPPLDFEALSQELGARDGTPPSPEDLLSAALYPKVYAEFRDFTSTFGPVSCGLTRLFLEG	1041
QY		1004	LVEGRETLIRLPDVRTPLLVRDLAISEPDDKGMNRNVVANVNGQIRPMRVDRSRVESVTAT	1063
Db		1042	PTIAEEFEVELERGKT-LHIKALALGDLNAAGOREAFFELNGQLRSILVTDQTQALKEMHV	1100
QY		1064	AEKADSSNKGHYAAPFAG-VVTVTVAEGDEVKAGDAVAII EAMKMEATTITASVDGKIDRV	1122
Db		1101	HPKADRSAKGQVGAPMPGEVWEVRVKEGEAVEBKAPLCVLSAMKMETVVVTAPREGGTVSRL	1160
QY		1123	VVPAATKVEGGDLIVVV	1139
Db		1161	HVRPGMSLEGGDLLIAEI	1177

PRELIMINARY;
 ID O7YS28
 RESULT 13
 Q7YS28
 PRT; 1178 AA.

RESULT 13
O7YS28

PRELIMINARY; PRT; 1178 AA.

AC Q7YS28;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pyruvate carboxylase.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae;
OX NCBI_TaxID=9823;

OS Suis scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC Suis scrofa (Pig).

carboxylase gene.";
 J. Anim. Breed. Genet. 0:0-0(2003).
 EMBL; AY225510; AAP57516.1; -.
 Pyruvate.
 SQ SEQUENCE 1178 AA; 129611 MW; F097ABF3D3124DA8 CRC64;

Query Match 43.9%; Score 2540.5; DB 6; Length 1178;
 Best Local Similarity 46.5%; Pred. No. 1:3e-126;
 Matches 544; Conservative 178; Mismatches 401; Indels 47; Gaps 13;

QY	6	SSTLPA-----FKKILVANRGEIAVRAAFRAALETGAATVAIYPREDRGSFHS	53
Db	19	TSTAPAAASPNNRRLVEYKPIKKVMVANRGEIAIRVFRACTELGIRTVAVYSEQDTGQMEHRQ	78
QY	54	FASEAVRIGTEGSPKAYLDIDEIIGAACKVKADAIYPGYGFLSENAQLARECAENGITF	113
Db	79	KADEAYLIGRLAPVOAYLHIPDIIKVAKENNVDAVHPGYGFLSERADFAQCQDAGVRF	138
QY	114	IGPTPEVLDLTGDKSRAVTAACKAGLPVL-AESTPSKNIDBIYKSAEGQTYPIFVKAVAG	172
Db	139	IGPSPVVRKMGDKVEARAIAAAGVPVPGTDAPITSLHEAHEFSNTYGFPIIFKAAAYG	198
QY	173	GGGRGMRFVASPDDELRLKLATEASREAEAAFGDGAVYVERAVINPOHIEVQILGDHTGEV	232
Db	199	GGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIKPRHIEVQILGDQYGNVL	258
QY	233	HLYERDCSLQRRHQKVEITAPAOHLDPELDRICADAVKFCRSIGYQGAGTVEFLVDEKG	292
Db	259	HLYERDCSIQRRHQKVELAPAAHLDPQLRTRLTSDSVKLAKQVYENAGTVEFLVDKHG	318
QY	293	NHVFIEMNPRIQVEHTVTEETVEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCR	352
Db	319	KHYFIEVNSRLQVEHTVTEETDVLVHAQIHVAEGRSLPDLGURQENIRINGCAIQCRV	378
QY	353	TTEDPNNGFRPDTGTITAYRSPCGAGVRLDGAQAOL-GGEITAHFDSMLVKMTCRGSDFET	411
Db	379	TTEDPARSFQPDGTGRIEVRSGEGMGIKLDNASAFQGAIVISPHYDSLIVKVIAHGKHPT	438
QY	412	AVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQG	471
Db	439	AATKMSRALAEFRVRGVKTNIPFLQNVLNQQLAGTVDTQFIDENBELFQLRPAQNRQAQ	498
QY	472	RILDYLAADVTVNKPVGVRP-KDVAAPIDK-LPNIKDLPLPRGSRDLKQLGPAAAFARDLR	529
Db	499	KLLHYLGHIMVNGPTTPIPVKANPSPDTPVVPVVPVPIGPPVAGFRDILLREGPEGFARAVR	558
QY	530	EQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVALKTPELLSVEAWGGATYDVAMRF	589
Db	559	NHQGLLLMDTTFRDAHQSLLATRVRTHDLKKISPYVAHNFSKLSIENWGGATFDVAMRF	618
QY	590	LPEDPWRLDELREAMPNVNIQMLLRGNTVGYTPYPDSVCRAFVKEAASSGVDFIRIFD	649
Db	619	LYECPWRLQLRELIPNIPFQMLLRGANAVGYTNPYDNVVFKECEVAKENGMDFVRVFD	678
QY	650	ALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDSDPNKLYTLTDYLLKMAEHIKSGAHI	709
Db	679	SLNVLPNLLGMEAVGSAG-GVVEAAISYTGVDADPSRTKYSLSQYYMDLABELVRAGTHI	737
QY	710	LAIKDMAGLLRPAAVTKLVTLRREF-DLPVHVHTHTTAGGQLATYFAAAQAGADAVDGA	768
Db	738	LCIKDMAGLLKPAACMLVSSLRDRFPDPLPHIHTDTSAGVAAMLACAQAGADVVDVA	797
QY	769	SAPLSGTTSQPSLSAIVAAFAHTRRTDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTG	828
Db	798	ADSMGNTSQPSLGALVACTRGTPDGTGVMEERVFYSEYWEAGRGLYAAFDCTATMKSG	857
QY	829	R--VYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGD	886
Db	858	NSDVYNEIIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGLDIKVTTPSSKIVGDL	917
QY	887	ALHLVGAGVDPAADFAADPKYDIPDSVIAFLRGLGNPPGGWPEPLRTRAL-----EGR	940
Db	918	AQFMVONGLSRAEAAQAEELSFPFRSVVEFLQYIGTTPHGGFPEPLRSKVLKDLPRVEGR	977

QY	941	SEGKAP-----LTEVPEEQAHLDADDKERRNSLNRLLEFPKPTTEEFLEHRRREF	990
Db	978	PGASLPDLQALEKELTERHGEVTPEDV-----LSAAMYPDVFAHFKDFTATFG	1028
QY	991	NTSALDDREFFYGLVEGRETLRLPDVRTPLLVRLDAISEPDDKGMNINVANVNGQIRPM	1050
Db	1029	PLDSLTRLFLQGPKEAEFEVELEERGKT-LHIKALAISDLNRAGORQVFFELNGQLRSI	1087
QY	1051	RVRDRSVESVTATAEKADSSNKGHVAAFPAG-VVTVTVAEGDEVKAGDAVAIIAMKMEA	1109
Db	1088	LVKDTQAMKEMHFHPKALKDVKGQIGAPMPGKVIDIKVAGAKVAKGQPLCVLSAMKMET	1147
QY	1110	TITASVDGKIDRVVVPAAATKVEGGDLIVVV	1139
Db	1148	VTSPMEGTVRKVHVTTDMTLEGGDDLILEI	1177

RESULT 14

Q819M9

ID	Q819M9	PRELIMINARY;	PRT;	1148	AA.
AC	Q819M9;				
DT	01-JUN-2003	(TREMBLrel. 24, Created)			
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DE		Pyruvate carboxylase (EC 6.4.1.1).			
GN	BC3947.				
OS	Bacillus cereus (strain ATCC 14579 / DSM 31).				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxID=226900;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=22608415; PubMed=12721630;				
RA	Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,				
RA	Kapatral V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,				
RA	Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,				
RA	Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,				
RA	Overbeek R., Kyrpides N.;				
RT	"Genome sequence of Bacillus cereus and comparative analysis with				
RT	Bacillus anthracis.";				
RL	Nature 423:87-91 (2003).				
DR	EMBL; AE017010; AAP10867.1; -.				
DR	GO; GO:0005737; C:cytoplasm; IEA.				
DR	GO; GO:0005524; F:ATP binding; IEA.				
DR	GO; GO:0016874; F:ligase activity; IEA.				
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.				
DR	GO; GO:0006094; P:gluconeogenesis; IEA.				
DR	GO; GO:0008152; P:metabolism; IEA.				
DR	InterPro; IPR005482; Biotin carb_C.				
DR	InterPro; IPR000089; Biotin_lipoyl.				
DR	InterPro; IPR005479; CPase_L_D2.				
DR	InterPro; IPR005481; CPase_L_N.				
DR	InterPro; IPR000891; HMGL-like.				
DR	InterPro; IPR003379; PYC_OADA.				
DR	InterPro; IPR005930; Pyruv carbbox.				
DR	Pfam; PF02785; Biotin carb_C; 1.				
DR	Pfam; PF00364; biotin_lipoyl; 1.				
DR	Pfam; PF00289; CPase_L_chain; 1.				
DR	Pfam; PF02786; CPase_L_D2; 1.				
DR	Pfam; PF00682; HMGL-like; 1.				
DR	Pfam; PF02436; PYC_OADA; 1.				
DR	TIGRFAMS; TIGR01235; pyruv carbbox; 1.				
DR	PROSITE; PS00866; CPASE_1; 1.				
DR	PROSITE; PS00867; CPASE_2; 1.				
KW	Ligase; Pyruvate; Complete proteome.				
SO	SEQUENCE 1148 AA; 1128442 MW; EE5CC5BA99D8E191 CRC64;				

Query Match 43.7%; Score 2528; DB 16; Length 1148;
Best Local Similarity 47.3%; Pred. NO. 5.6e-126;
Matches 543; Conservative 179; Mismatches 400; Indels 26; Gaps 13;

Db	4	LQRIQKVLVANRGEIAIRVFRACSELGLKTVIAIYSKEDSGSYHRYKADESYLVGEGKKPI	63
Qy	69	KAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS	128
Db	64	DAYLDIEGIIIEIAKSNHVDIAHPGYGFLSENIQFAKRCCEEGIIFIGPKSKHLDMGDKV	123
Qy	129	RAVTAAKKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRMRFVASPD	187
Db	124	KARTQAQLAQIPVPSGGPVNSLEEVEKFAEKVDYPIIIKASLGGGGRMIRVTS	183
Qy	188	RKLATEASREAAAFGAGVYVERAVINPOHIEVQILGDHTGEVVHLYERDCSLQRRHQ	247
Db	184	GESYNRAKSEAKAAGFNDVEYVEKFEKPKHIEVQILADEEGNVVHLYERDCSVQRRHQ	243
Qy	248	VVEIAPAQHLDPEDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEPNRIQVEH	307
Db	244	VVEIAPSVLSDDLQRICDAAVKLTKNVNYLNAGTVEFLVKD-DEFYFIEVNPVQVEH	302
Qy	308	TVTEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNN	364
Db	303	TITEMITGVDIVQSQIILADGSHLSKMGVGPKEEVAHVHGFIAIQSRVTEDPLN	362
Qy	365	TGFTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFEFAVARA	423
Db	363	TGKIMAYRSGGGFVRLDTGNSFGQAVITPYDLSLVKVTTWALTFEQAAKWERNLKEF	422
Qy	424	TVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLA	483
Db	423	RIRGIKTNIPLFENVVXKNFLSGEYDTSFIDASPELFLFPKRKDRGTKMLNYIGT	482
Qy	484	KPHGVRPKDVA-PIDKLPNIK-DLPLPRGSRDRKLQGPAAAFARDLREQDALAVT	541
Db	483	GFPGVKKEKPIFPDARIPCLKHSEPIONGTKQILDERGADGLVKWQDKRVLTT	542
Qy	542	RDAHQSLLATRVRSFALKPRAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDP	601
Db	543	RDAHQSLLATRIKDLHQIAEPTARMPLNLFSAEMWGATFDVAYRFLKEDPWER	602
Qy	602	REAMNVNIQMLLRGNTVGYTPYPDSVCRAFVKEAASSGVVDIFRIFDALNDVSQ	661
Db	603	REKMNVLQMLLRSSNAVGYKNYPDNLIQKFVECSAQAGIDVFRIFDSLNVWEG	662
Qy	662	DAVLETNTAFAEAMAYSGDLSDPNEKLYTLDDYLYLKMABEIVKSGAHILAIK	721
Db	663	DAVRDTG-KIAEATMCYTGDIDHPLRSKYDNLNYYKNLAKELEASGAHILG	721
Qy	722	AAVTKLVTALRREFDLPVHVHTDTAGGQLATYFAAAQAGADAVDGASAPLSGT	781
Db	722	NAAYDLVSALKETVSIPIHLHTDTSNGILTYTKAIEAGVDIVDAVSSMAGT	781
Qy	782	SAIVAAFAHTRRDTGLSLEAVSDLEPPYWEAVRGLYLPPESTGTPGTGRVYR	841
Db	782	NTLYYALGNGNERQPDVNIDSLEKLSHYWEDVRKYIAPPFSGMNAPTHTEV	841
Qy	842	NLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVYGDALHLVGA	901
Db	842	NLQQQAKAVGLGDRFDEVKVMYRRVNDMEGDIVKVTTPSSKVYGDMALEFM	901
Qy	902	ADPKYDIPDSVIAFLRGELGNPPGWPPEPLRTRALEGRSEBGKAPLTEVPEE	961
Db	902	ERGHAMDFPGSVVEMFSGDLGQPYGGFPKELQKIIL----KGKEPLTVRPG	957
Qy	962	DSKERR-NSLNR-----LLFPKPTBEEFLEHRRRRFGNTSALDDREFFYGL	1010
Db	958	ALKEELFKHLGREVTIFDVVAYALYPKVFMDEYKVAELYGNVSVLDTPTFF	1017
Qy	1011	LIRLPDVRTPLVRLDAISEPDDCKGMNVVANVNGQIRPMVRDRSRVESVTATA	1070
Db	1018	DVEIEQGT-LMVKLVSIGEPQPDGNNRVLYLEFNGQPPEIIVKDESVKATVA	1076
Qy	1071	NKGHVAAFPAG-VVTVTVTAEGDEVKAGDAVAIIEMKMEATTASVDGKIDR	1129
Db	1077	NPNHISATMPGTIVKVVVKEGDEVKKGDSMAITEAMKMETTVQAPFNGKVKV	1136

Qy	1130	VEGGDLIV	1137
	::	:	
Db	1137	IQTGDLII	1144
RESULT 15			
Q81MT6	PRELIMINARY; PRT; 1148 AA.		
ID	Q81MT6;		
AC	Q81MT6;		
DT	01-JUN-2003	(TREMBLrel. 24, Created)	
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Pyruvate carboxylase.		
GN	PYC OR BA4157.		
OS	Bacillus anthracis (strain Ames).		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=198094;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22608414; PubMed=12721629;		
RA	Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,		
RA	Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,		
RA	Holtzapfel E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,		
RA	Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,		
RA	DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,		
RA	Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,		
RA	Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,		
RA	Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,		
RA	Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,		
RA	Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,		
RA	Fraser C.M.;		
RT	"The genome sequence of Bacillus anthracis Ames and comparison to		
RT	closely related bacteria.";		
RL	Nature 423:81-86(2003).		
DR	EMBL; AE017037; AAP27881.1; --		
DR	TIGR; BA4157; --		
DR	GO; GO:0005737; C:cytoplasm; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0016874; F:ligase activity; IEA.		
DR	GO; GO:0004736; F:pyruvate carboxylase activity; IEA.		
DR	GO; GO:0006094; P:gluconeogenesis; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPR005482; Biotin_carb_C.		
DR	InterPro; IPR000089; Biotin_lipoyL.		
DR	InterPro; IPR005479; CPase_L_D2.		
DR	InterPro; IPR005481; CPase_L_N.		
DR	InterPro; IPR000891; HMGL-like.		
DR	InterPro; IPR003379; PYC_OADA.		
DR	InterPro; IPR005930; Pyruv_carbox.		
DR	Pfam; PF02785; Biotin_carb_C; 1.		
DR	Pfam; PF00364; biotin_lipoyL; 1.		
DR	Pfam; PF00289; CPsase_L_chain; 1.		
DR	Pfam; PF02786; CPsase_L_D2; 1.		
DR	Pfam; PF00682; HMGL-like; 1.		
DR	Pfam; PF02436; PYC_OADA; 1.		
DR	TIGRFAMS; TIGR01235; pyruv_carbox; 1.		
DR	PROSITE; PS00866; CPsase_1; 1.		
DR	PROSITE; PS00867; CPsase_2; 1.		
KW	Pyruvate; Complete proteome.		
SQ	SEQUENCE 1148 AA; 128573 MW; 57B97F8D9D1287BF CRC64;		

Query Match			
Best Local Similarity		43.6%;	Score 2526; DB 16; Length 1148;
Matches		542;	Pred. No. 7.1e-126;
		Conservative	178; Mismatches 402; Indels 26; Gaps 13;

Qy	9	LPAFKKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSPASEAVRIGTEGSPV	68
Db	4	LQRIQKVLVANRGEIAIRVFRACSELGLKTVIAIYSKEDSGSYHRYKADESYLVGEGKKPI	63
Qy	69	KAYLDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDLTGDKS	128
Db	64	DAYLDIEGIIIEIAKSNHVDIAHPGYGFLSENIQFAKRCCEEGIIFIGPKSKHLDMGDKV	123

Search completed: March 24, 2004, 22:44:15
Job time : 62 secs

QY	129	RAVTAAKKAGLPLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDEL	187
Db	124	KARTQAQLAQIPVPGSDGPVDSLEEVKEFAEKYDYP11IKASLGGGRMIRIVRTSEEL	183
QY	188	RKLATEASREAEAAFGDCAVYVERAVINPOHIEVQILGDHTGEVHLYERDCSLQRHOK	247
Db	184	RESYNRAKSEAKAAGNDEVYVEKFKPKHIEVQILADEEGNVVHLYERDCSVQRHOK	243
QY	248	VVEIAPAQHLDPFLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPRIOVEH	307
Db	244	VVEIAPSVLSDDLQRORICEAAVKLTKNVYNLNGTVEFLVKD-DNFYFIEVNPRVQVEH	302
QY	308	TVTEEVTEVDLVKAQMRLAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNGRPPD	364
Db	303	TITEMITGVDIVQSQIILADGHALHSQMVGPVKQEEVVVHGFALQSRVTTEDP1NNFMFD	362
QY	365	TGTITAYRSPGCGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDPFETAVARAQALAEF	423
Db	363	TGKIMAYRSGGFGVRLDTGNSFQGAVIAPYDSSLVKVTTWALTFEQAAAAMERNLKEF	422
QY	424	TVSGVATNIGFIRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLADEVN	483
Db	423	RIRGIKTNIPELENVVVKHKNFLSGEYDTSFIDASPELFLFKRKDRGTMKMLNYIGTVTN	482
QY	484	KPHGVRPKDVA-PIDKLPNI-KOLPLRGSRDRLKQLGPAAFARDLREQDALAVTDTTF	541
Db	483	GPPGVGKKEKPIPPDARIPNVLHSEPIQNGTKQILDERGADGLVKWVQDKRVLLTDTTF	542
QY	542	RDHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPPWDRDEL	601
Db	543	RDHQSLLATRIKTLKHQIAEPTARMPLNLFSAEMWGGATFDVAYRFLKEDPPWERLLDL	602
QY	602	REAMPNVNIQMLLEGRNTVGYTPYDPSVCRAFAVKEAASSGVDIPRIFDALNDVSQMRPAI	661
Db	603	REKMPNVLFQMLLRSSNAVGYKYPDNLIQKFVECSAQAGIDVFRIFDSLNWVEGMRVAI	662
QY	662	DAVLENTAVAEVAMAYSGLSDPNEKLYTLDYYLKMAEEIVKSGAHILAIKDMAGLLRP	721
Db	663	DAVRDTG-KIAEATMCYTGDIHDPMRSKYDLNYYKNLAKELEVSAGHILGIKDMAGLLKP	721
QY	722	AAVTKLVALTARREBFLPVHVHTHTAGGQATYFAAAQAGADAVDGASAPLSGTTSQPSL	781
Db	722	NAAYDLVSALKETVSIPIHLHTHTDTSNGILTYTKAIEAGVDIVDVAVSSMAGOTSQPSA	781
QY	782	SAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTGRVYRHEIPGGOLS	841
Db	782	NTLYYALGGNERQPDVNIDSLEKLSHYWEDVRKYVYAPFESGMNAPHTEVYMHMPGGQYS	841
QY	842	NLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTSSKVVGDALHLVGAGVDPADFA	901
Db	842	NLQQQAKVVGGLDRFDEVKVMYRRVNDMFGDIVKVTSSKVVGDMA1FMVQNHLTEQDVL	901
QY	902	ADPKYDIPDSVIAFLRGELGNPPGGMPEPLRTRALEGRSEBKAPLTEVPPEEQAHLDAD	961
Db	902	ERGHSMDFPGSVVEMFSGDLGQPYGGFPKQLQEIL----KGKEPLTVRPGELLEPVDPD	957
QY	962	DSKERR-NSLNR-----LLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRET	1010
Db	958	ALKEELFHKLGREVTMFDVVAYALYPKVFMDYEKVAELYGNVSVLDTPTFFYGMRLGEEI	1017
QY	1011	LIRLPDVRTPLLVRLDAISEPDDKGMNRNVVANVNGQIRPMRVDRDRSVESVTATAEKADSS	1070
Db	1018	DVEIEQGT-LMVKLVSIGELQPDGNRVLYLEFNGQPREIIVVKDESVKATVAQRVKGNRE	1076
QY	1071	NKGHVAAPFAG-VVTVTVVAGDEVKAGDAVAIEAMKMEATITASVDGKIDRVVVPAAATK	1129
Db	1077	NPNHISATMPGTVIKVVVKEGDEVKKGDSMAITEAMKMETTVQAPENGKVKVYVNDGDA	1136
QY	1130	VEGGDLIV	1137
Db	1137	IQTGDLLI	1144

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 24, 2004, 22:40:22 ; Search time 27 Seconds
(without alignments)
4061.420 Million cell updates/sec

Title: US-10-045-072-2
Perfect score: 5788
Sequence: 1 MSTHTSSTLPAPFKILVANR.....RVVVPAAATKVEGGDLIVVVS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3712.5	64.1	1127	2 D70671	pyruvate carboxyla
2	2541.5	43.9	1178	1 A47255	pyruvate carboxyla
3	2539.5	43.9	1178	2 JC4391	pyruvate carboxyla
4	2525.5	43.6	1150	2 A83978	pyruvate carboxyla
5	2524	43.6	1148	2 F69685	pyruvate carboxyla
6	2517.5	43.5	1178	1 JC2460	pyruvate carboxyla
7	2493	43.1	1146	2 AH1208	pyruvate carboxyla
8	2491.5	43.0	1144	2 D97227	pyruvate carboxyla
9	2490	43.0	1146	2 AC1565	pyruvate carboxyla
10	2488	43.0	1150	2 G89881	pyruvate carboxyla
11	2473.5	42.7	1174	2 AE2911	pyruvate carboxyla
12	2473.5	42.7	1174	2 C97686	pyruvate carboxyla
13	2464.5	42.6	1158	2 AE3285	pyruvate carboxyla
14	2464	42.6	1175	2 T20346	pyruvate carboxyla
15	2457	42.4	1178	1 QYBYP	pyruvate carboxyla
16	2443	42.2	1185	2 T39734	pyruvate carboxyla
17	2436	42.1	1180	2 S46094	pyruvate carboxyla
18	2419	41.8	1195	2 T43735	pyruvate carboxyla
19	2406	41.6	1137	2 E86708	pyruvate carboxyla
20	2115.5	36.5	984	2 T44608	pyruvate carboxyla
21	1079	18.6	501	2 D64453	biotin carboxylase
22	1063.5	18.4	477	2 G70427	biotin carboxylase
23	1036.5	17.9	447	2 A53311	biotin carboxylase
24	1036.5	17.9	447	2 AH1923	biotin carboxylase
25	1035	17.9	472	2 A70432	biotin carboxylase
26	1019	17.6	506	2 D69277	biotin carboxylase
27	997	17.2	491	2 A69123	biotin carboxylase
28	981.5	17.0	448	2 S74380	biotin carboxylase
29	976.5	16.9	471	2 G82966	probable biotin ca

30	975.5	16.9	1095	2 B83471	probable pyruvate
31	971.5	16.8	1078	2 D87647	hypothetical prote
32	961.5	16.6	447	2 B97338	biotin carboxylase
33	961	16.6	677	2 AC2997	hypothetical prote
34	959	16.6	667	2 F98286	hypothetical prote
35	957.5	16.5	444	2 C70444	biotin carboxylase
36	957	16.5	449	2 AI0912	biotin carboxylase
37	951.5	16.4	455	2 B86722	biotin carboxylase
38	949	16.4	449	2 AD0445	biotin carboxylase
39	948.5	16.4	539	2 T07093	acetyl-CoA carboxy
40	946	16.3	444	2 T44813	biotin carboxylase
41	946	16.3	448	1 F64105	biotin carboxylase
42	942.5	16.3	455	2 C95049	acetyl-CoA carboxy
43	942	16.3	450	2 A69581	acetyl-CoA carboxy
44	941.5	16.3	455	2 A97920	biotin carboxylase
45	939.5	16.2	457	2 H71553	probable biotin ca

ALIGNMENTS

RESULT 1

D70671
pyruvate carboxylase (EC 6.4.1.1) - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 11-Jan-2002
C;Accession: D70671; S73055
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genom
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70671
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1127 <COL>
A;Cross-references: GB:Z83018; GB:AL123456; NID:G3261671; PIDN:CAB05410.1; PID:G169486
A;Experimental source: strain H37RV
R;Smith, D.R.; Robison, K.
Submitted to the EMBL Data Library, September 1994
A;Description: Mycobacterium tuberculosis cosmid tbc2.
A;Reference number: S73053
A;Accession: S73055
A;Molecule type: DNA
A;Residues: 1-353; TRAGSARCDPPAVPVSAWTAAPTWRNQPVLRLHAGQADLS', 396-1115, 'EWRAETCWW' <S
A;Cross-references: EMBL:U00024; NID:G560506; PIDN:AAA50948.1; PID:G560527
C;Genetics:
A;Gene: pca; pyc
A;Start codon: GTC
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindin
C;Keywords: biotin binding; ligase; mitochondrion
F;4-457/Domain: biotin carboxylase homology <BCH>
F;1055-1127/Domain: lipoyl/biotin-binding homology <LPB>
F;1093/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 64.1%; Score 3712.5; DB 2; Length 1127;
Best Local Similarity 64.4%; Pred. No. 2.4e-196;
Matches 730; Conservative 153; Mismatches 239; Indels 11; Gaps 6;

QY	12	FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASAEVRIGTEGSPVKAY	71
DB	2	FSKVLVANRGEIAIRAFRAAVELGVGTVAVPYEDRNSQHLKADESYQIGDIGHPVHAY	61
QY	72	LDIDEIIGAARKVKADAIYPGCGFLSENAQLARECAENGITFIGTPEVLDLTGDKSAV	131
DB	62	LSVDEIVATARRAGADAIYPGCGFLSENPDLAACAAAGISFVGPSAEVLELAGNKSRAI	121
QY	132	TAACKAGLPVLAESTPSKNIDEIVKSAEGQTYPIFVKAVACGGGRGMRFVVASPDELKLA	191
DB	122	AAAREAGLPVLMSSAPSAASVDELLSVAAGMPFPLFVKAVACGGGRGMRRVGDIAALPEAI	181

QY 192 TEASREAAEAFDGAAYVERAVINPOHIEVQILGDHTGEVHLYERDCSLQRRHQKVEI 251
Db 182 EASREAEAFDGTPTVLEQAVINPRHIEVQILADNLGDVILHLYERDCSQRRHQKVIEL 241
QY 252 APAQHLDPQLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEHNPRIQVEHTVTE 311
Db 242 APAPHLDAELRYKMCVDAVAFARHIGYSCAGTVEFLDERGEYVFIEMNPRVQVEHTVTE 301
QY 312 EVTEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGPRPDTGTITAY 371
Db 302 EITDVLVASQLRIAAGETLEQLGRQEDIAHPGAALQCRITTEDPANGPRPDTGRISAL 361
QY 372 RSPGGAGVRLDGAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVATN 431
Db 362 RTAGGAGVRLDGSNLTGAEISFYFDSMLVKLTCRGRDLPTAVSRARRAIAEPRIRGVSTN 421
QY 432 IGFLRALLREEDFTSKRIATGRIADHPHLLQAPPADDEQGRILDYLDVTVNKPVGVRPK 491
Db 422 IPFLQAVLDDPDRAGRVTTSFIDERPQLLTARASADRGTKILNFLADVTVNNPYGSRPS 481
QY 492 DVAAPIDKLPNIKDLP-----PRGSRDLKQLGPAAFARDLREQDALAVDTTFRDAHQS 547
Db 482 TI-YPDDKLP--DLDLRAAPPAGSKQRLVKLGPFGFARWLRESAAVGVTDTTFRDAHQS 537
QY 548 LLATRVRSFALKPAAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRDLDELREAMPN 607
Db 538 LLATRVRTSGLSRVAPYLARTMPQLLSVECGGATYDVVALRFLKEDPWERLATLRAAMPN 597
QY 608 VNIQMLLRGRNTVGYTPYPDSVCRAFYKAEASSGVDFRIFDALNDVSMQRPDAIDAVLET 667
Db 598 ICLQMLLRGRNTVGYTPYPEIVTSAFVQEAATATGIDIFRIFDALNNIESMRPAIDAVRET 657
QY 668 NTAVAEVAMAYSGLSDPNKELYLDYYLYKMAEEIVKSGAHILAIDKMAGLLRPAAVTKL 727
Db 658 GSAIAEVAMCYTGDLTDPEGEQLYLDYYLYKLAEQIVDAGAHVLAIDKMAGLLRPPAAQRL 717
QY 728 VTALRREEDLPVHVHTHTDTAGGQLATYFAAAQAGADAVDAGASAPLSGTTSPQSLSAIVAA 787
Db 718 VSALRSREEDLPVHLHTHTDTTPGGQLASYVAAWHAGADAVDGAAPLAGTTSPALSSIVAA 777
QY 788 FAHTRRDGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIPGGQLSNLRAQA 847
Db 778 AAHTEYDTGLSLSAVCALEPYWEALRKVYAPFESGLPGPTGRVYHHEIPGGQLSNLRQQA 837
QY 848 TALGLADREELIEDNYAAVNEMLGPPTKVTPSSKVVGDLLALHLVGAGVDPADPAADPQKY 907
Db 838 IALGLGDRFEEIEEAYAGADRVLGRVLKVTPSTSKVVGDLLALALVGAGVSADEFASDPARF 897
QY 908 DIPDSVIAFLRGELGNPPGWPPEPLRTRALEGRSEKAPLITEVPEEEQAHLDADDSKERR 967
Db 898 GIPESVTLGRLGELGDPGGWPPEPLRTAALAGGAAR-PTAQLAADDEIALSSVGAK-RQ 955
QY 968 NSLNRLLFPKPTKEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDA 1027
Db 956 ATLNRLLPSPPTKEFNEHREAYGDTSQLSANQFFYGLRQGBEHRVKL-ERGVELLIGLEA 1014
QY 1028 ISEPDCKMRNVVANVNGQIRPMRVDRDSVESVTATAEKADSSNKGHVAAFPAGVVTVTV 1087
Db 1015 ISEPDERGMRTVMCILNGQLRPVLVRDRSIAAVPAAEKADRGNPGHIAAPFAGVVTVG 1074
QY 1088 AEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVVPAATKVEGGDLIVVVS 1140
Db 1075 CVGERVAGQGTIATIEAMKMEAPITAPVAGTVERVAVSDTAQVEGGDLLVVVS 1127

RESULT 2
A47255
C; pyruvate carboxylase (EC 6.4.1.1) precursor [similarity] - mouse
C; Species: Mus musculus (house mouse)
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Feb-2002
C; Accession: A47255
R; Zhang, J.; Xia, W.L.; Brew, K.; Ahmad, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 1766-1770, 1993

A; Title: Adipose pyruvate carboxylase: amino acid sequence and domain structure deduced
A; Reference number: A47255; MUID:93189578; PMID:8446588
A; Accession: A47255
A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-1178 <ZHA>
A; Cross-references: GB:L09192; NID:G293743; PIDN:AAA39737.1; PID:G293744
A; Experimental source: 3T3-L1 adipocytes
A; Note: sequence extracted from NCBI backbone (NCBIN:126874, NCBIPI:126875)
C; Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C; Keywords: biotin binding; ligase; mitochondrion
F; 1-20/Domain: transit peptide (mitochondrion) #status predicted <TRP>
F; 21-1178/Product: pyruvate carboxylase #status predicted <MAT>
F; 39-494/Domain: biotin carboxylase homology <BCH>
F; 1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F; 1144/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 43.9%; Score 2541.5; DB 1; Length 1178;
Best Local Similarity 47.0%; Pred. No. 7.4e-132;
Matches 537; Conservative 187; Mismatches 401; Indels 17; Gaps 11;

QY 13 KKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSFASEAVRIGTEGSPVKAYL 72
Db 38 KKMVANRGEIAIRVFRACTELGI RTVAVYSEQDTGQWHRQKADEAYLIGRGLAPVQAYL 97
QY 73 DIDEIIGA AKKVKADAIYPGYGLSENQAALARECAENGITFIPTPEVLDLTGDKSRAVT 132
Db 98 HIPDIK VAKENGVDVHPGYGLSERADFAQCQDAGVRFIGSPSEVVRKMGDKVEARA 157
QY 133 AAKKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVASPDELKLA 191
Db 158 IAIAGVPVPGTDSPISSLHEAHEFSNTFGFPIIFKAAAYGGGRGMRVHVHSELEENY 217
QY 192 TEASREAAEAFDGAAYVERAVINPOHIEVQILGDHTGEVHLYERDCSLQRRHQKVEI 251
Db 218 TRAYSEALAAFGNGALFVEKIEKPRHIEVQILGDQYGNILHLYERDCSIQRHQQKVEI 277
QY 252 APAQHLDPQLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEHNPRIQVEHTVTE 311
Db 278 APATHLDPQLRSRLTSDSVKLAKQVYENAGTVEFLVDKHKHYFIEVNSRLQVEHTVTE 337
QY 312 EVTEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGPRPDTGTITAY 371
Db 338 EITDVLVASQLRIAAGETLEQLGRQEDIAHPGAALQCRITTEDPANGPRPDTGRISAL 397
QY 372 RSPGGAGVRLDGAAQL-GGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVAT 430
Db 398 RSGEGMGI RLDNASAFQGA VISP HYDSL LVKVIAHGKDHTAATKMSRALAEFRVGKVT 457
QY 431 NIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDVTVNKPVGVRP 490
Db 458 NIPFLQVNLNNQQFLAGTVDTQFIDENPEL FQLRPAQNRAQKLLHLYLGHVMVNGTTPIP 517
QY 491 KDVA-APID-KLPNIKDLP LPRGSRDLRLKQLGPAAFARDLREQDALAVDTTFRDAHQS 548
Db 518 VNVSPSPVDPVAVPVPIGPPAGPRDILLREGPEGFAPAVRNHQGLLLMDTTFRDAHQS 577
QY 549 LATRVRSFALKPAAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWDRDLDELREAMPN 608
Db 578 LATRVRTDHLKKIAPYVAHNFNKLFSMENWGGATFDVAMRFLYECWRRRLQELRELIPNI 637
QY 609 NIQMLLRGRNTVGYTPYPDSVCRAFYKAEASSGVDFRIFDALNDVSMQRPDAIDAVLETN 668
Db 638 PFQMLLRGANAVGYTNPVDPNVVFKFCEVAKENGMDVFRVFDLSNLYLPNMLLGMEAGSAG 697
QY 669 TAVAEVAMAYSGLSDPNKELYLDYYLYKMAEEIVKSGAHILAIDKMAGLLRPAAVTKLV 728
Db 698 -GVVEAAISYTG DVADPSRTKISLEYMGLAEELVRAGTHILC IKDMAGLLKPAACTMLV 756
QY 729 TALRREF-DLPVHVHTHTDTAGGQLATYFAAAQAGADAVDAGASAPLSGTTSPQSLSAIVAA 787
Db 757 SSLRDRFPDLP LHIHTHTDTSGAGVAAMLACACAQAGADVVDVAVDMSMGMTSPQSMGALVAC 816

QY 788 FAHTRDRTGLSLEAVSDLEPYWEAVRGILYLPFFESGTPGPTGR--VYRHEIPGGQLSNLR 845
Db 817 TKGTPLDTEVPLERVDYSEYWEAGRLYAADFCTATMKSGNSDVYENEIPGGQYTNLHF 876
QY 846 QATALGLADRFELIEDNYAAVNEMLCRPTKVTTPSSKVVGDLALHLVGAGVDPADFAADPQ 905
Db 877 QAHSMGLGSKPKVKKAYVEANQMLGDLIKVTPSSKIVGDLAQFMVQNGLSRAEAEQA 936
QY 906 KYDIPDSVTAPLRGELGNPPGGWPELRTAL-----EGRSEGKAPLTEVPPEEQAHLD 959
Db 937 ELSFPRSVEFLQGYIGIPHGFPPEPERSKVLKDLPRIEGRPGASLPPLNLKELEKDLID 996
QY 960 A-DDSKERRNSNRLRLLFPKPTTEFLEHRRRFGNTSALDDREFFYGLVEGRETLRLDVR 1018
Db 997 RHGEVTPEDVLSAAMYPDVFAQKDFATFGPLDSLNLRLFLQGPKEAEFEVELERGK 1056
QY 1019 TPLLRLDAISEPDDKGMNRNVANVNGQIRPMRVRDRSVESVTATAEKADSSNKGHVAAP 1078
Db 1057 T-LHIKALAVSBLNRAGORQVFFELNGQLRSILVKDTQAMKEMHFPKALKDVKGQIGAP 1115
QY 1079 FAG-VVTVTVAEGDEVKAGDAVAIIIEAMKWEATITASVDGKIDRVVVVPAATKVEGGDLIV 1137
Db 1116 MPGKVIDIKVAAGDKVAKGQPLCVLSAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLIL 1175
QY 1138 VW 1139
Db 1176 EI 1177
RESULT 3
JC4391
pyruvate carboxylase (EC 6.4.1.1) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jan-1996 #sequence revision 25-Apr-1997 #text_change 11-Jan-2002
C;Accession: S68252; S72393; JC4391; S06440
R;Jitrapakdee, S.; Booker, G.W.; Cassady, A.I.; Wallace, J.C.
Biochem. J. 316, 631-637, 1996
A;Title: Cloning, sequencing and expression of rat liver pyruvate carboxylase.
A;Reference number: S68252; MUID:96257760; PMID:8687410
A;Accession: S68252
A;Molecule type: mRNA
A;Residues: 1-1178 <JIT1>
A;Cross-references: EMBL:U36585; NID:g1040973; PIDN:AAC52668.1; PID:g1040974
A;Accession: S72393
A;Molecule type: protein
A;Residues: 489-505 <JIT2>
A;Experimental source: liver
R;Lehn, D.A.; Moran, S.M.; MacDonald, M.J.
Gene 165, 331-332, 1995
A;Title: The sequence of the rat pyruvate carboxylase-encoding cDNA.
A;Reference number: JC4391; MUID:96096548; PMID:8522203
A;Accession: JC4391
A;Molecule type: mRNA
A;Residues: 1-221, 'P', 223-865, 'D', 867-976, 'G', 978-1178 <LEH>
A;Cross-references: GB:U32314; NID:g929987; PIDN:AAA96256.1; PID:g929988
A;Experimental source: liver
R;Thampy, K.G.; Huang, W.Y.; Wakil, S.J.
Arch. Biochem. Biophys. 266, 270-276, 1988
A;Title: A rapid purification method for rat liver pyruvate carboxylase and amino acid s
A;Reference number: S06440; MUID:89024676; PMID:3178228
A;Accession: S06440
A;Molecule type: protein
A;Residues: 'SG', 23-25, 'PL', 28-29, 'LL', 32-34, 'P', 1134, 'A', 1136-1137, 1139-1152, 'T', 1154-1
C;Comment: This enzyme is located in the mitochondrial matrix and catalyzes the conversi
intermediates that exit the mitochondrion for consumption in various pathways.
C;Genetics:
A;Gene: pc
A;Genome: nuclear
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: biotin binding; gluconeogenesis; homotetramer; ligase; mitochondrion
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;21-1178/Product: pyruvate carboxylase #status predicted <MAT>
F;39-494/Domain: biotin carboxylase homology <8CH>

F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (Lys) (covalent) #status experimental
Query Match 43.9%; Score 2539.5; DB 2; Length 1178;
Best Local Similarity 46.5%; Pred. No. 9.5e-132;
Matches 542; Conservative 185; Mismatches 410; Indels 29; Gaps 12;
QY 1 MSTHTSSTLPA-----FKKILVANRGEIAVRAALETGAATVAIYPREDRG 48
Db 14 LGVRRSSTAPVASPNVRRLEYKPIKKVMVANRGEIAIRVRACTELGIRTVAVYSEQDTG 73
QY 49 SFHRSFASEAVRIGTEGSPVKAYLDEIIGAAKKVADAIYPGYGFLSENAQLARECAE 108
Db 74 QMHRQKADAEAYLIGRLAPVQAYLHIPDIIKVAKENGVDVHPGYGFLSERADFAQACD 133
QY 109 NGITFIGPTPEVLDLTGDKSRVTAACKAGLPVL-AESTPSKNIDEIVKSAEQTYPIFV 167
Db 134 AGVRFIGPSPEVVRKMGDKVEARALAAAGVPVPGTNSPINSLHEAHEFSNTYGFPIIF 193
QY 168 KAVAGGGGRGMRFVASPDELRLKLA TEASREAEAAFGDGVYVERAVINPQHIEVQILGDH 227
Db 194 KAAVGGGGRGMRVHVHSYELEENYTRAYSEALAAFGNGALFVEKFKIEKPRHIEVQILGDQ 253
QY 228 TGEVHLYERDCSLQRRHQKVEIAPAQHLDPELRDRICADAVKPCRISIGYQAGTVEFL 287
Db 254 YGNILHLYERDCSLQRRHQKVEIAPATHLDPQLRSRLTSDSVKLAKQVGYENAGTVEFL 313
QY 288 VDEKGNHVFIEPNRIQVEHTVTEEVTEVDLVKAQMRLAAGATLKLGLTQDKIKTHGAA 347
Db 314 VDKHGKHVFIENSRQLQVEHTVTEETDVLVHAQIHVSEGRSLPDLGLRQENIRINGCA 373
QY 348 LQCRITTEDPNNGPRPDTGITAYRSPGGAGVRLDGAAL-GGEITAHFDSMLVMTCRG 406
Db 374 IQCRVTTEDPARSPQPDGTGRIEVRSGEGMIRLDNASAFQGAIVISPHYDSLVLVKVIAHG 433
QY 407 SDFETAVARAQALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPA 466
Db 434 KDHPATAATKMSRALAEFRVGVKTNIPFLQNLNNQQFLAGIVDTQFIDENPELFLQLRPA 493
QY 467 DDEQGRILDYLDVTVNKPVGVRPKDVA-APIDKL-PNIKOLPLPRGSRDRLKQLGPAAF 524
Db 494 QNRAQKLLHLYLGHVMVNGPTTPIPVKVSPPVDPIVPVPIGPPAGFRDILLREGPEGF 553
QY 525 ARDLREQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYD 584
Db 554 ARAVRNHQGLLLMDTTFRDAHQSLLATRVTHDLKKIAPYVAHNFNFLFSIENWGGATFD 613
QY 585 VAMRFLFEDPDWRLDELREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFAVKEAASSGVDI 644
Db 614 VAMRFLYECFWRRLQELRELIPNIPFQMLLRGANAVGYTNYPDNVVFKEVAKENGMDV 673
QY 645 FRIFDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGLDSDPNEKLYTLDYYLKMAGEEIVK 704
Db 674 FRIFDSLNYLPNMLLGMEAAGSAG-GVVEAAISYTGADVADPSRTKYSLSEYVMGLAEELVR 732
QY 705 SGAHILAIKDMAGLLRPAAVTKLVTLRREF-DLPVHVHTHTDTAGGQLATYFAAAQAGAD 763
Db 733 AGTHILCIKDMAGLLKPAACTMLVSSLRDRFPDLPLHIHTHTDTSGSGVAAMLACAQAGAD 792
QY 764 AVDGASAPLSGTTSPSLSAI VAAFAHTRRRDTGLSLEAVSDLEPYWEAVRGILYLPFESGT 823
Db 793 VVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVDYSEYWEAGRLYAADFCTA 852
QY 824 PGPTGR--VYRHEIPGGQLSNLRQAQATALGLADRFELIEDNYAAVNEMLCRPTKVTTPSSK 881
Db 853 TMKSGNSDVYENEIPGGQYTNLHFQAHSMGLGSKPKVKKAYVEANQMLGDLIKVTPSSK 912
QY 882 VVGDALHLVGAGVDPADFAADPKYDIPDSVTAPLRGELGNPPGGWPELRTAL----- 937
Db 913 IVGDLAQFMVQNGLSRAEAEQAEELSFPRSVMVVEFLQGYIGIPHGFPPEPERSKVLKDL 972
QY 938 --EGRSEGKAPLTEVPPEEQAHLDA-DDSKERRNSNRLRLLFPKPTTEEFLEHRRRFGNTSA 994

Db 973 RIEGRPGASLPPLNLEKLEKDLIDRHGHEEVPEDVLSAAMPYDVFAQKDFATATGFLDS 1032

QY 995 LDDREFFYGLVEGETLIRLPDVRTPLVRLDAISEPDDKGMNVVANVNGQIEPRVRD 1054

Db 1033 LNTRLFLOQPKIAEFEEVELEKGT-LHIKALAVSDLNRAQORQVFFELNGQLRSILVKD 1091

QY 1055 RSVESVTATAEKADSSNKGHVAAFPAG-VVTVTVAEAGDEVKAGDAVAIIEMKMEATITA 1113

Db 1092 TQAMKEMHFPKALKDVKQIGAPMPGKVIDVKVAGAKVVKQPLCVLSAMKMETVTVTS 1151

QY 1114 SVDGKIDRVVVPAAKVEGGDLIVV 1139

Db 1152 PMEGTIRKVHVTKDMTLEGDDLILEI 1177

RESULT 4

A83978

pyruvate carboxylase pyca [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: A83978

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; PMID:20512582; PMID:11058132

A;Accession: A83978

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1150 <STO>

A;Cross-references: GB:AP001516; GB:BA000004; NID:G10175192; PIDN:BA06344.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: pyca

C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 43.6%; Score 2525.5; DB 2; Length 1150;

Best Local Similarity 46.8%; Pred. No. 5.4e-131;

Matches 539; Conservative 185; Mismatches 400; Indels 27; Gaps 13;

QY 9 LPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSEAFSEAVRIGTEGSPV 68

Db 4 LKNIKKVLVANRGEIAIRIFRACTELHRTVAIYKEDTGAHYRYKADEAYLVGEGKKPI 63

QY 69 KAYLDIDEIIIGAAKKVKADAIYPGYGFLSENAQLARECAENGITFISGPTPEVLDTGDKS 128

Db 64 EAYLDIEGIIIEIAKHGVDAIHYPGYGFLSENIIEFAKRCHCEGIFIGPELEHLVMEFGDKV 123

QY 129 RAVTAACKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRVFASPDLE 187

Db 124 QAREQAIAKANLPVPGSDGCVSSLEDVKAPADKHGYFFIIKALGGGRGMRIVRSENDV 183

QY 188 RKLATEASREAAAFDGAIVYVERAVINPOHIEVQILGDHTGEVHLRYERDCSLQRRHQK 247

Db 184 QESYERAKSEAKAAGFNDVYVEKFIENPKHIEVQILADKHGNTLHLYERDCSVQRRHQK 243

QY 248 VVEIAPAQHLDPEDLRICADAVKFCRSIGYQAGTVFELVDEKGNHVFIEMNPRIOVEH 307

Db 244 VVEVAPSVLSLSEDRERICQAAVQLAENNVYVNAVGTVEFLVDRGNFYFIEVNPRIQVEH 303

QY 308 TVTEEVTEVLDLVKAQVRLAAGATL--KEGL-TQDKIKTHGAALQCRITTEDPNNRPRPD 364

Db 304 TITEMVTGIDIVQSLFIADGEHLHGDRGLGIPKQEEIVCHGYATQSRVTTEDPSNGFLPD 363

QY 365 TGTITAYRSPGGAGVRILD-GAAQLGGEITAHPDSMLVKMTCRGSDFETAVARAQALAEF 423

Db 364 TGRINAYRSGGGFVRLDAGNGFQGAIVTPYDLSLVKSVTWTALTFEGAAKMLNREF 423

QY 424 TVSGVATNIGFLRALREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDVTVN 483

Db 424 RIRGIKTNIAFLENVVRHQFSLSGEYNTSFIDQTPFELVFPKRDGRTKMLSFIGETIVN 483

QY 484 KPHGVRPKDVAAPIDK--LPNIK-DLPLPRGSRDLKQLGPAAPAFARDLREQDALAVTDTT 540

Db 484 GYPGLE-KTKKPVFDKPPVPKLLKSEPIPDGKTQILDQHGPEGLAKWVKEQKHVLLTDTT 542

QY 541 PRDAHQSLLATRVSPALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFLFEDPWRILDE 600

Db 543 PRDAHQSLLATRVTHDLKQIAEPTARLLPNLFSMEMWGGATFDVAMRFLHEDPWERILLI 602

QY 601 LREAMPNVIQMLLRGRNTVGYTPYDSVCRAFVKEAASSGVDFIRIFDALNDVSVQWRPA 660

Db 603 LRKCAPNVLFOMLLRASNAVGYKNYPDNLIREFVDKSNAGIDVFRIFDSLNVWVEGMKLA 662

QY 661 IDAVLETNATAVAEAMAYSGDLSDPNEKLYTLDDYLLKMAEEIVKSGAHILAIKDMAGLLR 720

Db 663 IEAVGEAN-KIAEATICYTGDIIDSSRPKYDLAYYKLLAKELEAAGAHILGKDMAGLLK 721

QY 721 PAAVTKLVTALRRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVDGAASPLSGTTSOPS 780

Db 722 PEAAVQLVAELKDTVTIPVHLHTHTDTSNGIFTYARAIEAGVIDVDVAVSSMAGLTSQPS 781

QY 781 LSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESETPGPTGRVYRHEIPGGQL 840

Db 782 ANSLYALADSERQPNVNITALEQAEFEWETRKFYAGFESGWNAPHTVEYHEMPGGQY 841

QY 841 SNLRAQATALGLADRFEIEDNYAAVNEMLGRPTKVTPSSKVVGDLALHLVAGVDPADF 900

Db 842 SNLQQAQAKAVGLGHRWNEVKMYRTVNDMFGDVVKVTPSSKVVGDMALYMQNDLTEBEV 901

QY 901 AADPQKYDIPDSVIAFLRGLGNPPGGWPEPLRTRALEGRSECKAPLTEVPEEEQAHLDA 960

Db 902 YENGKLDLDFPDSVVEFFEGQLGQPYQGPKKLEIILKGRK----PITNRPGENMEPIQF 957

QY 961 DDSKER-RNSLNR-----LLFPKPTTEEFLEHRRRFGNTSALDDREFFYGLVEGRE 1009

Db 958 EAIKEELYNKLDROVTSHDILSYALYPKVFMEFERFRQTFEGVSLDTPTFYGLRPGEE 1017

QY 1010 TLIRLPDVRTPLVRLDAISEPDDKGMNVVANVNGQIRPMRVRDRSVESVTATAEKADS 1069

Db 1018 IVEIEQKGT-LIVKFISSLKPDQDGNRIYVYFELNGQPREVLKQSVKTSIISRPKADK 1076

QY 1070 SNKGHVAAFPAG-VVTVTVAEAGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVVPAA 1128

Db 1077 SNPNHIGASMPGTVVVKALVEKDKVKQGDHLMITEAMKMETTVQAPFDGEVVALHVKDGD 1136

QY 1129 KVEGGDLIVV 1139

Db 1137 AIQTGDLLEIV 1147

RESULT 5

F69685

pyruvate carboxylase (EC 6.4.1.1) pyca [similarity] - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000

C:Accession: F69685

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Bruchet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetelli

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: F69685

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-1148 <KUN>

A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13359.1; PID:g2633857
A;Experimental source: strain 168
C;Genetics:
A;Gene: pyca
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding
C;Keywords: ligase
F;8-465/Domain: biotin carboxylase homology <BCH>
F;1073-1146/Domain: lipoyl/biotin-binding homology <LPB>
F;1112/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 43.6%; Score 2524; DB 2; Length 1148;
Best local similarity 46.0%; Pred. No. 6.5e-131;
Matches 533; Conservative 186; Mismatches 392; Indels 48; Gaps 13;

QY 11 AFKKILVANRGEIAVRAFAALETGAATVAIYPRDRGSHRSPASEAVRIGTEGSPVKA 70
DB 5 SIQKVLVANRGEIAIRIFRACTELNIRTVAVYSKEDSGSYHRYKADEAYLVGEGKPIDA 64

QY 71 YLDIDEIIGAARKVKADAIYPGYGLSENAQLARECAENGITFIGPTPEVLDLTGDKSRA 130
DB 65 YLDIEGIIDIAKRNKVDIHPGYGLSENIHFARRCEEIGIVFGPKSEHLDMFGDKVKA 124

QY 131 VTAACKAGLPVLAEST-PSKNIDEIVKSABEQTYPIFVKAVAGGGGRGMRFVASPDELK 189
DB 125 REQAEKAGIPVPGSDGPAETLEAVEQFGQANGYPIIKASLGSGGRGMRIVRSESEVKE 184

QY 190 LATEASREAEAAFGDGAIVYVERAVINPOHTEVQILGDHTGEVHLHYERDCSLQRRHQKV 249
DB 185 AYERAKSEAKAAGFNDVYVEKLIENPKHIEVQVIGDKQGNVHLFERDCSVQRRHQKVI 244

QY 250 EIAPAQHLDPRLDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPRIQVEHTV 309
DB 245 EVAPSVLSPELRDQICEAAVALAKNVYINAGTVEFLV-ANNEFFYFIEVNPVQVEHTI 303

QY 310 TEEVTEVDLVKAQMRLAAGATL--KELGLTQDK-IKTHGAALQCRITTEDPNNGFERPDTG 366
DB 304 TEMITGVDIVQTIILVAQGHSLHKKVNIPEQKDIFTIGYAIQSRVTTEDPQNDMPDTG 363

QY 367 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFFETAVARAQALAEFTV 425
DB 364 KIMAYRSGGGFGVRLDTGNSFQAVITPYDYSLLVXLSTWALTFFEQAAAKVNRNLQEFRI 423

QY 426 SGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLAADVTVN-- 483
DB 424 RGIKTNIPELVNAKHEKFLTGQYDTSFIDITPELFNFPKQKDRGTGKMLTYIGNVTVNGF 483

QY 484 -----KPHGVVRKDVAAPIDKLPNIKDLPLPGSRDRLLKQLGPAAPARDLREQDALAV 536
DB 484 PGIGKKEKPAFKPLGVKVDVDQPP-----ARGTKQILDEKGAEGLANVVKQKSVLL 536

QY 537 TDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLPEDPWD 596
DB 537 TDTTFRDAHQSLLATRIKSHDLKKIANPTAALWPELFSMEMWGGATFDVAYRFLKEDPWK 596

QY 597 RLDELREAMPNVNIQMLLRGNTVGYTPYPSVCRAFVKEAASSGVDFRIFDALNDVSQ 656
DB 597 RLEDLRKEVPNTLFQMLLRSSNAVGYTNPYDNVKEFVKQSAQSGIDVFRIFDSLNVWVG 656

QY 657 MRPAIDAVLENTAVAEVAMAYSGDSDPNEXKLYTLDYILKMAEEIVKSAHILAUKDMA 716
DB 657 MTLAIDAVRDTG-KVAEAAICYTGDILDKNRTKYDLAYYTSMAKELEAAGAHILGIKDMA 715

QY 717 GLLRPAAVTKLVALTALRREFDLPHVHTHTDAGGQALATYFAAAQAGADAVDGASAPLSGTT 776
DB 716 GLLKPAAYELVSALKETIDIPVHLHTDTSNGIYMYAKAVEAGVDIIDVAVSSMAGLT 775

QY 777 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 836
DB 776 SQPSASGFYHAMEGNDRRPEMNVQGVVELLSQYWESVRKYSEFESGMKSPHTEIYHEMP 835

QY 837 GGQSLNRAQATALGLADRFEFLIEDNYAAVNEMLGRPTKVTPTSSKVVGDALHLVAGAVD 896
DB 836 GGQYENLQQQAKGVGLGDRMNEVKEMYRRVNDMFGDIVKVTPTSSKVVGDMLVMVQNNLT 895

QY 897 PADFAADPPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGKAPLTEVPDEE--- 953
DB 896 EKDVYKGESLDFPDSVVVELPKGNIGQPHGGFPEKLQKILKQGE-----PITVRPGEELLE 951

QY 954 -----EQAHLDAADSKERRNSLNRLLPKPTBEEFLEHRRRFGNTSALDDREFF 1001
DB 952 PVSFEAIKQEFKEQHNLEISD-----QDAVAYALYPKVFTDYVKTTESYGDISVLDTPFF 1007

QY 1002 YGLVEGRETLIRLPDVRTPLLVLDALISEPDDKGMNVVANVNGQIRPMRVDRDRSVESVT 1061
DB 1008 YGMTLGEIEVEIERGKT-LIVKLISIGEPOPDA TRVVYFELNGQPREVVIKDESIKSSV 1066

QY 1062 ATAEKADSSNKGHVAAPFAGVVTVTVAE-GDEVKAGDAVAIIEAMKMEATITASVDGKID 1120
DB 1067 QERLKADRTNPISHIAASMPGTVIKVLAEAGTKVKNKGHDHLMINEAMKMETTVOAPFSGTIK 1126

QY 1121 RVVVPAAATKVEGGDLIVVV 1139
DB 1127 QVHVKNGEPIQTGDLLEI 1145

RESULT 6
JC2460
pyruvate carboxylase (EC 6.4.1.1) precursor - human
N;Alternate names: pyruvate:carbon dioxide ligase (ADP-forming)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence revision 18-Aug-2000 #text_change 01-Feb-2002
C;Accession: G01933; JC2460; B27883; S01469
R;Walker, M.E.; Jitrapakdee, S.; Val, D.L.; Wallace, J.C.
submitted to the EMBL Data Library, July 1995
A;Reference number: H00708
A;Accession: G01933
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1178 <WAL>
A;Cross-references: EMBL:U30891; NID:g1101028; PIDN:AAA82937.1; PID:g1101029
R;MacKay, N.; Rigat, B.; Douglas, C.; Chen, H.S.; Robinson, B.H.
Biochem. Biophys. Res. Commun. 202, 1009-1014, 1994
A;Title: cDNA cloning of human kidney pyruvate carboxylase.
A;Reference number: JC2460; MUID:94324922; PMID:8048912
A;Accession: JC2460
A;Molecule type: mRNA
A;Residues: 1-224, 'WP', 227-351, 'A', 353-384, 'PT', 387-485, 'DV', 488-637, 'R', 639-728, 'A', 7
A;Cross-references: GB:S72370; NID:g632807; PIDN:AAB31500.1; PID:g632808
R;Lamhonwah, A.M.; Quan, F.; Gravel, R.A.
Arch. Biochem. Biophys. 254, 631-636, 1987
A;Title: Sequence homology around the biotin-binding site of human propionyl-CoA carbo
A;Reference number: A27883; MUID:87212051; PMID:3555348
A;Accession: B27883
A;Molecule type: mRNA
A;Residues: 1083-1178 <LAM>
A;Cross-references: GB:M26122; NID:g189657; PIDN:AAA36423.1; PID:g387003
R;Freytag, S.O.; Collier, K.J.
J. Biol. Chem. 259, 12831-12837, 1984
A;Title: Molecular cloning of a cDNA for human pyruvate carboxylase. Structural relati
A;Reference number: S01469; MUID:85030380; PMID:6548474
A;Accession: S01469
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1135-1178 <FRE>
A;Cross-references: EMBL:K02282; NID:g189655; PIDN:AAA60033.1; PID:g189656
C;Genetics:
A;Gene: GDB:PC
A;Cross-references: GDB:119472; OMIM:266150
A;Map position: 11q11-11q13.1
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindin
C;Keywords: biotin binding; gluconeogenesis; ligase; mitochondrion
F;1-20/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;21-1178/Product: pyruvate carboxylase #status predicted <MAT>
F;39-494/Domain: biotin carboxylase homology <BCH>
F;1105-1178/Domain: lipoyl/biotin-binding homology <LPB>
F;1144/Binding site: biotin (Lys) (covalent) #status predicted

Query Match	43.5%;	Score 2517.5;	DB 1;	Length 1178;
Best Local Similarity	46.2%;	Pred. No. 1.5e-130;		
Matches 536;	Conservative 184;	Mismatches 412;	Indels 29;	Gaps 12;
QY 6	SSTLPA-----FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSEFHS	53		
DB 19	TSTAPASPNNRRLEYKPIKKVMVANRGEIAIRVFRACTELGIRTVAIYSEQDTGMHRQ	78		
QY 54	FASEAVRIGTEGSPVKAYLDIDEIIGAAGKVKADAIYPGYFLSENAQLARECAENGITF	113		
DB 79	KADEAYLIGRLAPVQAYLHIPDIIKVAKENNVDAVHPGYFLSERADFAQACQDAGVRF	138		
QY 114	IGTPEVLDTGDKSRAVTAACKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAG	172		
DB 139	IGPSPEVVRKMGDKVEARAIAAAGVPVPDGTAPITSLHEAHEFSNTYGFPIIFKAAAYG	198		
QY 173	GGGRGMRFVASEDELRLKLATEASREAEAFGDGAVYVERAVINPQHIEVQILGDHTGEVV	232		
DB 199	GGGRGMRVVHSEYELEENYTRAYSEALAAFNGALFVEKFIKPRHIEVQILGDQYGNIL	258		
QY 233	HLYERDCSLQRHQKQVVEIAPAQHLDPQLRDLICADAVKFCRSIGYQAGTVEFLVDEKG	292		
DB 259	HLYERDCSIQRHQKQVVEIAPAAHLDPQLRDLTSDSVKLAKQVYENAGTVEFLVDRHG	318		
QY 293	NHVFIEPNRIQVEHTVTEEVTEVDLVKAQVRLAAGATLKEGLTQDKIKTHGAALQCRI	352		
DB 319	KHYFIEVNSRLQVEHTVTEEITDVLVHAQIHVSEGRSLPDLGLRQENIRINGCAIQCRV	378		
QY 353	TTEDPNNGFRPDTGTTAYRSPGGAGVRLDGAQL-GGEITAHFDSMLVKMTCRGSDFET	411		
DB 379	TTEDPARSFQDGTGRIEVRSGEGMIRLDNASAFQGAVISPHYDSSLVKVIAHGKDHT	438		
QY 412	AVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLOAPPADDEQG	471		
DB 439	AATKMSRALAEFRVRGVKTNIAFLQNLNNQFLAGTVDTQFIDENPELFQLRPAQNRAQ	498		
QY 472	RILDYLAADVTKNPKHGVRP-KDVAAPIDK-LPNIKDLPLPRGSRDLKQLGPAAPFARDLR	529		
DB 499	KLLHYLGHVWVGPTTPIPVKASPSPTDPVPAVPTGPPAGFRDILLREGPEGFARAVR	558		
QY 530	EQDALAVDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRF	589		
DB 559	NHPGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLFMSMENWGGATFDVAMRF	618		
QY 590	LFEDPWDRLDELREAMPNVNIQMLLRNTVGYTPYDPSVCRAFAVKEAASSGVDFRIED	649		
DB 619	LYECPWRRLOELRELI PNIPFQMLLRGANAVGYTNPVNVVFKFEVAKENGMDVFRVFD	678		
QY 650	ALNDVSOMRPAIDAVLETNTAVAEVAMAYSGDLSDBNEKLYTLDDYLLKMAEEIVKSGAHI	709		
DB 679	SLNYLPNMLLGMEAGSAG-GVVEAAISYTGVDADPSRTKYSLOXYMGLAEELVRAGTHI	737		
QY 710	LAIKDMAGLLRPAAVTKLVTALRREF-DLPVHVHTHTAGGQLATYFAAAQAGADAVDGA	768		
DB 738	LCIKDMAGLLKPTACTMLVSSLRDRFPDPLPLHIHTHTSGAGVAAMLACAQAGADVDDVA	797		
QY 769	SAPLSGTTSQPSLSAIVAAFAHTRRDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTG	828		
DB 798	ADSMGMSGTSQPSMGALVACTRGTPDTEVPMERVFYSEYEWEGARGLYAAAFDCTATMKSG	857		
QY 829	R-VYRHEIPGGQSLNRAQATAGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLL	886		
DB 858	NSDVYENEIPGGQYTNLHFQAHSMGLGSKFKEVKKAYVEANQMLGDLLIKVTTPSSKIVGDLL	917		
QY 887	ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPEGGWPEPLRTRAL-----EGR	940		
DB 918	AQFMVQNGLSRAEAEAAQAEELSFPFRSVVEFLQYIGVPHGCGFPFPRFSKVLKDLPRVEGR	977		
QY 941	SEKCAPLTEVPEEEQAHLDA-DDSKERRNSLNRLLFPKPTKEEFLEHRRRRFGNTSALDDRE	999		
DB 978	PGASLPPLDQALEKELVDRHCEEVTPEDVLSAAMYDPVFAHFKDFTATFGPLDSLNTRL	1037		

QY 1000 FFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMRNVTNVANVNGQIRPMRVDRSRVES 1059
| | : | : | : | : | : | : | : | : | : | :
Db 1038 FLOGPKIAEFEVELERGKT-LHIKALAVSDLNLRAGQRQVFPELNGQLRSILVKDTQAMK 1096
| | : | : | : | : | : | : | : | : | : | :
QY 1060 VTATAEKADSSNKGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIEAKMKEATITASVDGK 1118
| | : | : | : | : | : | : | : | : | : | :
Db 1097 EMHFHPKALKDVKGQIGAPMPGKVIDIKVAGAKVAKGGQPLCVLSAMKMETVVTSPMEGT 1156
| | : | : | : | : | : | : | : | : | : | :
QY 1119 IDRVVVPAATKVEGGDLIVV 1139
: : | : | : | : | : | :
Db 1157 VRKVHVTKDMTLEGDDLILEI 1177
| | : | : | : | : | : | : | : | : | : |

RESULT 7
AH1208
pyruvate carboxylase homolog pyca [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AH1208
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.: Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1208
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1146 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99150.1; PID:g16410474; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: pyca
C;Superfamily: pyruvate carboxylase: biotin carboxylase homology; lipoyl/biotin-binding


```
QY 488 VRPKDVAAPIDKLNIXDLP-----LPRGSRDLKQLGPAAFARDLRBDALAVDTDTFRD 543
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
484 IKHRD--KPVYAEPLPKIPYGSQISPGTKILDAKGPVGVDVWKQEEVLLTDTTLRD 541
QY 544 AHQSLLATRVRSFALKPAAEAVALKTPELLSVEAWGGATYDVAMRFLFDPDRLDELRE 603
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
542 AHQSLLATRVRSKDIPOVADAMAHLLPNMFSEFMWGGATFDVAYRFLNEDPWRLTLRK 601
QY 604 AVPNVNIQMLLRGNTVGYTPYDPSVCRAFYKEAASSGVDIFRIFDALNDVVSQMRPAIDA 663
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
602 QIPVMFQMLLRGANAVGKYNPDNVIREFVKSAQSAGVDVFRVPSLWIKGMEVSIDA 661
QY 664 VLETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKVAEEIVKSGAHILAIDKMACGLLRPAA 723
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
662 VREAG-KVVEAAICTGDIDDDTRTKYTTIDYKDMAKELVAQGTILGIDKMACGLLKPOA 720
QY 724 VTKLVTLRRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDAGASAPLSGTTSPQSLSA 783
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
721 AYRLIGELKDTVDVPIHLHHTDTSNGGIYTYAAAVSAGVDIVDVASSAMSGATSPSMTG 780
QY 784 IVAAPAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFPESGTPGPTGRVYRHEIPIGGQLSNL 843
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
781 LYYGLVNGNRQTNLDAQNSQIINHYYWEDVRHYKDFDNALNSPQTEVIYIHEMPGGQYTNL 840
QY 844 RAQATALGLADRFELIEDNYAAVNEMLRPTKVTTPSSKVVGDLALHLVAGVDPADFAAD 903
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
841 QQQAIAVGLGDRWDEVKEMYTVVNQMFGDIVKVTTPSSKVVGDLALFMVQNELSEEDVYEK 900
QY 904 PQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEEEAHLDDADS 963
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
901 GDTIDFPDSVIEFFMGBIGQPYGGFPEKQLVLKGR-----PLTDRPGALMEPVNFVDV 956
QY 964 K-----ERRNSLNRLLPFKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLI 1012
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
957 KAELEKMGYEPTKDVISYILYPKFVLDYQDMINKYGDVTLDTPTPKGMRLGETIEV 1016
QY 1013 RLPDVRTPLLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEKADSSNK 1072
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1017 ELEKGKI-LLIKLNSIGEPDAGTRVIYFELNGQPREINIQDMNVQSTVIARRKIDTTNP 1075
QY 1073 GHVAAPFAG-VVTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKIDRVVVVPAATKVE 1131
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1076 EHVGMATMTGSVIVQVVVKGDSVKKGDPDLLITEAMKMETTIQAPFDGEVSSIYVSDGDTIE 1135
QY 1132 GGDLIWVS 1140
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1136 SGDLLIEVN 1144

RESULT 8
D97227
pyruvate carboxylase, PYKA [imported] - Clostridium acetobutylicum
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C/Accession: D97227
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97227
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1144 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80607.1; PID:g15025689; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2660
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 43.0%; Score 2491.5; DB 2; Length 1144;
Best Local Similarity 45.3%; Pred. No. 4e-129;
Matches 520; Conservative 204; Mismatches 391; Indels 33; Gaps 14;
```

```
QY 12 FKXILVANRGEIIVRAFRAALETGAATVAIYPREDRGSEHRSFASEAVRIGTEGSPVKAY 71
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 FXRVLVANRGEIIVRAFRACHLGIIRTVAIYSEEDKLALFRTKADESILIGNQKGPVDAY 64
QY 72 LDIDEIIGAANKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEVLDLTGDKSRV 131
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 LNIIDEIINLALKKGVDAIHPGYGFLSENSEFSRRCTEAGIEFIGTGMMDKLGDKINSK 124
QY 132 TAAKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFRVSPDELRLK 190
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 LAAKAAAGVKTIIPGVEKPIETEQAIEFARTCGYPVMVAAAAGGGGRGMRIVEKEEDLIAA 184
QY 191 ATEASREAAEAFGDAVYVERAVINPQHIEVQILGDHTGEVVHLVERDCSLQRRHQKVE 250
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 CRSAKSEAKKAFGIEDIFIEKYLEGPXKHIEVQVLGDYKGNIVHLVERDCSVQRRHQKVIE 244
QY 251 IAPAQHLDPEDRLDRIICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEHNPRIQVEHTVT 310
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 LTPAVSMSEEKRLCEICEDALKIARSIGYRSAGTLEFLLDKGNHYFIEMNPRVQVEHTIT 304
QY 311 EEVTEVDLVKAQMLAAGATLK--ELGL-TQDKIKTHGAALQCRITTEDPNNGFRPDGTGT 367
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
305 EMVTGIDIVQSILLIAEGYKLNPSVEGINSQEDIHVNGYAIQCRITTEDPSNSFAPDTGK 364
QY 368 ITAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVS 426
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
365 IDVYRTSGSGFIRLDGNGGFTGAVISPYDYLKSTWSRTFEDAIRKAIRAIKETIYIS 424
QY 427 GVATNIGFLRALLREEDFTSKRIATGTGFIADPHLLQAPPADDEQGRILDYLDLVTVNKPX 486
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
425 GVKTNIDFLINVLNHEHTRFKGLCDTNFIANNPFLFEITPRIDTELRLVKFIEGKVVNETH 484
QY 487 GVRPKDVAAPIDKLP--NIKDLPLRGSRRDLKQLGPAAFARDLREQDALAVTDTTFRDA 544
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
485 GHK---IEVDVPSVPKYEIKE-PL-RGTKQILDEKPGKGLVEWIKDQDKLLTDTTMRDA 539
QY 545 HQSLLATRVRSFALKPAAEAVALKTPELLSVEAWGGATYDVAMRFLFDPDRLDELREA 604
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
540 HQSLMATRLRTVDMVKIAKAESVLAKDLFSMEMWGGATFDTAYRFLKESPWERLERLKR 599
QY 605 MPNVNIQMLLRGNTVGYTPYDPSVCRAFYKEAASSGVDIFRIFDALNDVVSQMRPAIDAV 664
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
600 VPNVLFQMLLRGANAVGKYNYPDNVIREFIKQSSKSGIDVFRIFDSLNVWKGMEVAIDEV 659
QY 665 LETNTAVAEVAMAYSGDLSDPNEKLYTLDYILKVAEEIVKSGAHILAIDKMACGLLRPAAV 724
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
660 LNQG-KVAEACMCYTGDLTDNRDKYTLNYYVNLAKIEKSGAHILGIDKMSALLKPYAA 718
QY 725 TKLVTLRRREFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDAGASAPLSGTTSPQSLSAI 784
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
719 LKLIRALKNEISIPILHHTDTTGNGVATVLMAAHAGVDIADTAFNSMSGSLTSQPALNSV 778
QY 785 VAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFPESGTPGPTGRVYRHEIPIGGQLSNLR 844
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
779 VAAALKNTDRDTKMDIGDLQKISDYWSTVRPVYKFSGLKAVSAEIIKYEIPGGQYSLNK 838
QY 845 AQATALGLADRFELIEDNYAAVNEMLRPTKVTTPSSKVVGDLALHLVAGVDPADFAADP 904
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
839 PQVESFGLGHRFEQVKEMYREVNIMLGDIVKVTTPSSKVVGDLAIFMVQNELTSENILEKA 898
QY 905 QKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEVPEEEEAHLDDADS 964
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
899 KDMFFPDSVVSFYKGMGMQPKGPGPKELQKIVL----KDEEAITCRPGELLDPDEDKIR 954
QY 965 ERNNSLNRL-----LFPKPTEEFLEHRRRFGNTSALDDREFFYGLVEGRETLIR 1013
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
955 VRLUKENKIEPTDKDVISYALYPDVFEYLYKYKNEYGDLSRMGSDVFFHGLAEGEISELE 1014
QY 1014 LPVYRTPLLVRLDAISEPDDKGMNVVANVNGQIRPMRVDRSVESVTATAEK---ADSS 1070
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1015 IAEGKT-LWVQLLHIGLKDQGNRTLVFEVNGNRRRRIKIKDKVSTKSEIVEIVEIADSS 1073
```


Db 5 KKLIVANRGEIAIRFRAAAELDISTVAIYSNEDKSSLHRYKADESYLVGSDLGPAESYL 64

Qy 73 DIDEIIGAAKVKADAIYPGYGFLSENAQLARECAENGITFIGTPPEVLDLTGDKSRVAT 132

Db 65 NIERIIDVAKQANVDAIHPGYGFLSENEQFARRCAEAGIKFIGPHEHLDHFGDKVKART 124

Qy 133 AAKKAGLPVL-AESTPSKNIDEIVKSAEGQYPIFVKAVAGGGGRGMRFAVSPDELRLKLA 191

Db 125 TAIKADLPVIPGTGPIKSYELAKEFAEEAGFPLMIKATSGGGGKGMRIVREESELEDAF 184

Qy 192 TEASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLORRHQKVVEI 251

Db 185 HRAKSEAEKSGNSEVIERIDNPKHIEVQIGDEHGNIVHLFERDCSVORRHQKVVEV 244

Qy 252 APAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLV--DEKGNHVFIEMLNPRIOVEHTV 309

Db 245 APSVGLSPTLRQRICDAAIQLMENIKYVYAGTVEFLVSGDE---FFFIEVNPVQVEHTI 301

Qy 310 TEEVTEVDLVKAQMRLAAGATL~-KELGLTQDK-IKTHGAALQCRITTEDPNNNGFRPDTG 366

Db 302 TEMVTGIDIVKTQILVAAGADLFGEIINPQOKDITTLGYAIQCRITTEDPLNDFMPDTG 361

Qy 367 TITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALAEFTV 425

Db 362 TITAYRSSGGFVRLDAGDGFQGAIEISPYDSELVVKLSTHAISFKQABEKMVRSLEMRI 421

Qy 426 SGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDADVTVNKP 485

Db 422 RGVKTNIPLINVMKNKFTSGDYTTKFIETPELFDIOPSLDRGTKTLEYIGNVTIN-- 479

Qy 486 HGVPRKQVAAPIDKLPNIKDLPLP-----RGRDRLKQLGPAAFARD 527

Db 480 -----GFPNVEKRPKPDYELASIPTVSSSKIASFSGTKQLLDEVGPKGVAEW 526

Qy 528 LREQDALAVTDTTFRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAM 587

Db 527 VKQDDVLLTDTTFRDAHQSLLATRVRTKMINIASKTADVFKOGFSLEWGGATFDVAY 586

Qy 588 RPLFEDPWDRLDLREAMPNVNIQMLLRGNTVGYTPYPDSVCRAVFEAAASSGVDIFRI 647

Db 587 NFLKENPWERLERLKAIPNVLFQMLLRASNAVGYKNYPDNVIHKFVQESAKAGIDVFI 646

Qy 648 FDALNDVSQMRPAIDAVLETNTAVAEVAMAYSGDLSDP-NEKLYTLDYVLKMAEEIVKSG 706

Db 647 FDSLNVVDQMKVANEAVQAG-KISEGTICYTGDIILPERSNIYTTYLEYVVKLAKELEREG 705

Qy 707 AHILAIDMAGLLRPAAVTKLVTALRREFDLPVHVHTHTAGGQLATYFAAAQAGADAVD 766

Db 706 FHILAIDMAGLLKPKAAVELIGELKSAVDLPILHHTHTDSGNLLTYKQAI DAGVDIID 765

Qy 767 GASAPLSGTSQPSLSAIVAAFAHTRRDRTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGP 826

Db 766 TAVASMSGLTSQPSANSLYALNGFPRLHRTDIEGWSLSHYWSTVRTYYSDFESDIKSP 825

Qy 827 TGRVYRHEIPGGQLSNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDL 886

Db 826 NTEIYQHEMPGGQYSNLSQQAQSLGLGERFDEVKDMYRRVNFELGDIVKVTTPSSKVVGDM 885

Qy 887 ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAP 946

Db 886 ALYMQNDLDEQSVITDGYKLDFFESVVSFFKGEIGQPVNGFNKDLQAVILKQGE---A 941

Qy 947 LTEVPEEEQAHLDDDSKE-----RRNSLNLLFPKPTTEEFLEHRRRFGNTSAL 995

Db 942 LTARGEYLEPVDFEKVRLELLEEEQQGPVTEQDIISVLYPKVYEQIYQIOTRQYGNLSLL 1001

Qy 996 DDEFFYGLVEGRETLIRLPVTRPPLLVRDLAISEPDDKGMNVVANVANGQIRPMVRDR 1055

Db 1002 DTPTEFFGMNGETVEIEI-DXGKRLIIKLETISEPDENGNRTIYYAMNGQARRIYIKDE 1060

Qy 1056 SVESVTATAEKADSSNKGHVAPFAGVVT-VTVAEQDEVKAGDAVAIIEMKMEATITAS 1114

Db 1061 NVHTNANVKPKADKSNPNSHIGAQMFGSVTEVKSVGETVKANQPLLITEAMKMETTIQAP 1120

Qy 1115 VDGKIDRVVVPAAATKVEGGDLIVVV 1139

Db 1121 FDGVIKQVTVNNGDTIATGDLLEI 1145

RESULT 11

AE2911

pyruvate carboxylase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C;Accession: AE2911

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, R;Wood, D.W.; Gallet, W.; Grant, C.; Guenthrner, D.; Kutayavin, T.; Levy, R.; Li, M.; McCleerage, G.; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AE2911

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1174 <KUR>

A;Cross-references: GB:AE008688; PIDN:AAL43707.1; PID:g17741236; GSPDB:GN00186

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: pyca

A;Map position: circular chromosome

C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 42.7%; Score 2473.5; DB 2; Length 1174;

Best Local Similarity 46.6%; Pred. No. 4.1e-128;

Matches 541; Conservative 173; Mismatches 406; Indels 41; Gaps 19;

Qy 8 TLPAPFKILVANRGEIAVRAFAALETGAATVAIYPREDRGSHRSFASAEAVRIG----- 62

Db 20 TVLKISKILVANRSEIAIRVFRAANELGIKTVAIWAEEDKLSLHRFKADESIVQVGRPHL 79

Qy 63 -TEGSPVKAYLDIDEIIGAAKVKADAIYPGYGFLSENAQLARECAENGITFIGTPPEVL 121

Db 80 AKDMGPIESYLSIEEVIRVAKLSGADAIHPGYGLLSSESPEFVEACNKAGITFIGTPDTM 139

Qy 122 DLTGDKSRAVTAAKKAGLPVLAESTP-SKNIDEIVKSAEGQYPIFVKAVAGGGGRGMRF 180

Db 140 RQLGNKVAARNLAISVDVPVVPATNPLPDDIAEVERMAEEIGYPVMLKASWGSGGRGMRA 199

Qy 181 VASDDELRLKLA TEASREAAAFGDGAVYVERAVINPQHIEVQILGDHTGEVVHLYERDCS 240

Db 200 IRKXEDLAREVTEAKREAAAFGKDEVYLEKLVERRARHVESQILGDTHGNVVHLFERDCS 259

Qy 241 LQRHQKVVEIAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVD-EKGNHVFIEML 299

Db 260 IQRNQKVVERAPAPYLSAQEQELAAYSLSKIAAATNYIGAGTVEYLMADATGKFEFIEV 319

Qy 300 NPRIQVEHTVTEEVTEVDLVKAQMRLAAGATL~-KELGL-TQDKIKTHGAALQCRITTED 356

Db 320 NPRIQVEHTVTEVTGIDIVKAQIHILEGAAIGTASGVPKQEDIRLNGHALQCRITTED 379

Qy 357 PNNGFRPDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVAR 415

Db 380 PEHNFIPDYGRITAYRSASGFGIRLDGGTSYTGAVITRYDPLLVKVTAWAPEPEDAISR 439

Qy 416 AQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILD 475

Db 440 MDRALREFRIRGVATNLTFLBAIIGHDSFRNNTYTTFRIDSTPELFQAQVKRQDRATKLLT 499

Qy 476 YLADVTVN---KPHG-VRPKDVAAPIDKLPNIKDLPLPRGSRDRKLQLGPAAFARDLRE 530

Db 500 YLADVTVNGHPETKGRAPKPADKAAKPI--VPYI-DAPTEDDGTQQLLDKLGPKGFADWMRN 556

Qy 531 QDALAVTDTTFRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMRFL 590

Db 557 EKRVLVTDITMRDGHQSLLATRVTRTHDIARVASVSKALPQLLSLECGGATFDVSMRFL 616
QY 591 FEDPWRDLDELREAMPNVNIQMLLRGRNTVGYTYPDSCVRAEFVKEAAASSGVDFIRIFDA 650
Db 617 TEDPWERLSLIREGAPNLLQMLLRGANGVGYKNYPDNVVKYFVRQAARGVDLFRVFD 676
QY 651 LNDVSQMRPAIDAVLENTAVAEVAMAYSGLSDPNKLYTLDYLLKMAEEIVKSGAHIL 710
Db 677 LNWVENMRVSMDAIAEEN-KLCEATICYTGDLNLSARPKYDLKYTNLAVELEKAGAHII 735
QY 711 AIKDMAGLLRPAAVTKLVTLRREFFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASA 770
Db 736 AVKDMAGLLKPAKAAKVLFKALREATGLPIHFHTHTDTSISAAATVLAADVDAVDAAMD 795
QY 771 PLSGITTSQPSLSAIVAAFAHTRRTDGLSLEAVSDLEPYNEAVRGLYLPFESGTPGPTGRV 830
Db 796 AFSGNTSQPCLGSIVEALSISRDTGLDTEWIRRIISFYWEAVRNQYAAFESDLKGPASEV 855
QY 831 YRHEIPGGQSLNLAQATAGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDALHL 890
Db 856 YLHEMPGGQFTNLKEQARSIGLESRWHEVAQAYADANRMFGDIVKVTTPSSKVVGDWALMM 915
QY 891 VGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEV 950
Db 916 VSQDLTVADVENPDREVSPDSVVSMLKGDLGQSPGGWPEALQKKAL----KGEKPYTVR 971
QY 951 PEE--EQAHLDDADDSKERRNSLNR-----LLFPKPTEEFLEHRRRRFGNTSALDDR 998
Db 972 PGSILLEADLDA-ERKVIETKLERKVDDDFEFASYLMYPKVFTDFALTAEYGVSVLP 1030
QY 999 EFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMNVNANVNGQIRPMRVRDRS-V 1057
Db 1031 AYFYGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFFFEINGQPRRIKVPDRAHG 1089
QY 1058 ESVTATAEKADSSNKGHVAAFPAGVVT-VTVAEGDEVKAGDAVAIEAMKMEATITASVD 1116
Db 1090 ASGSAVRRKAEPCGNASHIGAPMPGVISRVRFINQGVKAGDVLLSIEAMKMETALHAERD 1149
QY 1117 GKIDRVVVPAAATKVEGGDLIV 1137
Db 1150 GKIAEVLVKPGQIDAKDILLI 1170

RESULT 12
C97686
pyruvate carboxylase (U51439) [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: C97686
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: C97686
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1174 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK88444.1; PID:g15157941; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_4940
A;Map position: circular chromosome
C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

Query Match 42.7%; Score 2473.5; DB 2; Length 1174;
Best Local Similarity 46.6%; Pred. No. 4.le-128;
Matches 541; Conservative 173; Mismatches 406; Indels 41; Gaps 19;
QY 8 TLPAFKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVRIG----- 62
Db 20 TVLKISKILVANRSEIAIRVFAANELGIKTVAIWAEDKLSLHREFKADESQVGRGPHL 79
QY 63 -TEGSPVKAYLDIDEIIGAAGKVKADAIYPGYGFLSENAQLARECAENGITFIGTPEVL 121

Db 80 AKDMGPIESYLSIEEVIRVAKLSGADAIHPGYLLSESPEFVEACNKAGITFIGTPTDM 139
QY 122 DITGDKSRAVTAAKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMR 180
Db 140 RQIGNKVAARNLAISVDVPVVPATNPLPDDIAEVRMAEEIGYPVMLKASWGSGGRMRA 199
QY 181 VASPDDELRLKLAATEASREAEAFGDGAVYVERAVINPQHIEVQILGDHTGEVVLIERDCS 240
Db 200 IRKKEDLAREVTEAKREAAAFGKDEVYLEKLEVERARHVESQILGDTHGNVHLFERDCS 259
QY 241 LQRHQKVVEIAPAQHLDPBLDRICADAVKFCRSIGYQAGTVEFLVD-EKGNHVFIE 299
Db 260 IQRRNQKVVVERAPAPYLSEARQELAAVSLKIAATNYIGAGTVEYLMADATGKPYFIEV 319
QY 300 NPRIQVEHTVTEEVTEVDLVKAQMRLLAAGATL--KELGL-TQDKIKTHGAALQCRITTED 356
Db 320 NPRIQVEHTVTEVTGIDIVKAQIHILEGAAICTAESGVPKQEDIRLNGHALQCRITTED 379
QY 357 PNNGFERPDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDEFETAVAR 415
Db 380 PEHNFIDYGRITAYRSASGFGIRLDGGTSYTGAVITRYDDYDPLLVKVTAWAPEDEAISR 439
QY 416 AQRALAEFTVSGVATNIGFIRALLREEDFTSKRIATCFIADHPHLLQAPPADDEQGRILD 475
Db 440 MDRALREFRIRGVATNLTFLBAIGHDSFRNNTYITTFIDSTPELFAQVKRQDRATKLLT 499
QY 476 YLADVTVN---KPHG-VRPKDVAA-PIDKLPNIKDLPLPRGSRDRLLKQLGPAAPAFDLRE 530
Db 500 YLADVTVNGHPETKGRAPKADKAAKPI--VPYI-DAPTDDGTQQLLDKLGPKGFADWNRN 556
QY 531 QDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFL 590
Db 557 EKRVLVTDITMRDGHQSLLATRVTRTHDIARVASVSKALPQLLSLECGGATFDVSMRFL 616
QY 591 FEDPWRDLDELREAMPNVNIQMLLRGRNTVGYTYPDSCVRAEFVKEAAASSGVDFIRIFDA 650
Db 617 TEDPWERLSLIREGAPNLLQMLLRGANGVGYKNYPDNVVKYFVRQAARGVDLFRVFD 676
QY 651 LNDVSQMRPAIDAVLENTAVAEVAMAYSGLSDPNKLYTLDYLLKMAEEIVKSGAHIL 710
Db 677 LNWVENMRVSMDAIAEEN-KLCEATICYTGDLNLSARPKYDLKYTNLAVELEKAGAHII 735
QY 711 AIKDMAGLLRPAAVTKLVTLRREFFDLPVHVHTHDTAGGQLATYFAAAQAGADAVDGASA 770
Db 736 AVKDMAGLLKPAKAAKVLFKALREATGLPIHFHTHTDTSISAAATVLAADVDAVDAAMD 795
QY 771 PLSGITTSQPSLSAIVAAFAHTRRTDGLSLEAVSDLEPYNEAVRGLYLPFESGTPGPTGRV 830
Db 796 AFSGNTSQPCLGSIVEALSISRDTGLDTEWIRRIISFYWEAVRNQYAAFESDLKGPASEV 855
QY 831 YRHEIPGGQSLNLAQATAGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDALHL 890
Db 856 YLHEMPGGQFTNLKEQARSIGLESRWHEVAQAYADANRMFGDIVKVTTPSSKVVGDWALMM 915
QY 891 VGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEKAPLTEV 950
Db 916 VSQDLTVADVENPDREVSPDSVVSMLKGDLGQSPGGWPEALQKKAL----KGEKPYTVR 971
QY 951 PEE--EQAHLDDADDSKERRNSLNR-----LLFPKPTEEFLEHRRRRFGNTSALDDR 998
Db 972 PGSILLEADLDA-ERKVIETKLERKVDDDFEFASYLMYPKVFTDFALTAEYGVSVLP 1030
QY 999 EFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGMNVNANVNGQIRPMRVRDRS-V 1057
Db 1031 AYFYGMEDGEELFADIERGKTLVIVN-QASSGIDDKGMVTVFFFEINGQPRRIKVPDRAHG 1089
QY 1058 ESVTATAEKADSSNKGHVAAFPAGVVT-VTVAEGDEVKAGDAVAIEAMKMEATITASVD 1116
Db 1090 ASGSAVRRKAEPCGNASHIGAPMPGVISRVRFINQGVKAGDVLLSIEAMKMETALHAERD 1149
QY 1117 GKIDRVVVPAAATKVEGGDLIV 1137

Db 1150 GKIAEVLVXPGDQIDAKDLLI 1170

RESULT 13

AE3285

pyruvate carboxylase (EC 6.4.1.1) [imported] - Brucella melitensis (strain 16M)

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002

C;Accession: AE3285

R;Delvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: AE3285

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1158 <KUR>

A;Cross-references: GB:AE008917; PIDN:AAL51448.1; PID:g17982157; GSPDB:GN00190

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEI0266

A;Map position: 1

C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-binding

C;Keywords: ligase

Query Match 42.6%; Score 2464.5; DB 2; Length 1158;

Best Local Similarity 46.6%; Pred. No. 1.2e-127;

Matches 538; Conservative 171; Mismatches 408; Indels 37; Gaps 15;

QY 14 KILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVRIG-----TEGSP 67

Db 10 KILVANRSEIAIRVFAAANELGLKTVTIWAEEEDKLSLHRFKADESQVGRGPHLDRDLGP 69

QY 68 VKAYLDIDEIIIGAAKVKADAIYPGYGFLSENAQLARECAENGITFCPTPEVLDLTGDK 127

Db 70 IESYLSIDEIIRVAKLSGADAIHPGYGLLSEPFSAEACAENGIVFCGPKPETMRRLLGNK 129

QY 128 SRAVTAAKKAGLPVLAESTP-SKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFAVSPDE 186

Db 130 VAARNLAIEIGVPVVPATDPLPDDMDEVKLAAGIYPLMLKASWGGGGRGMRAIRAEAD 189

QY 187 LRKLATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRRHQ 246

Db 190 LAREVMEAKREAKAAGFQKDEYVLEKLVERRARHVEVQILGDTYGNVHLFERDCSIQRRNQ 249

QY 247 KVVETAPQAHLDPEDLRICADAVKFCRSIGYQAGTVEFLVD-EKGNHVFIEMNPRIOV 305

Db 250 KVVERAPAPYLNDQRRRELADYGLKIAHATDYIGACTVEFLMDADTGKFFIEVNPRIQV 309

QY 306 EHTVTEEVTEVDLVKAQMRILAAGATL--KELGL-TQDKIKTHGAALQCRITTEDPNNNGFR 362

Db 310 EHTVTEEVTGIDIVKQIHILIEGFAIGTPESGVPRQEDIRLNGHALQCRITTEDPEQNFI 369

QY 363 PDTGTITAYRSPGGAGVRLD-GAAQLGGEITAHFDSMLVKMTCRGSDFETAVARAQRALA 421

Db 370 PDYGRIQAYRSAAGFIRLDGTAYSGAFITRYVDPLLVKVTASGATPLEATHRMDRALR 429

QY 422 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDVDT 481

Db 430 EFRIRGVATNLTFLFAIINHFKPLSNDYTTFRIDTTPPELFEQMKRQDRATKLLTYIADVT 489

QY 482 VN-----KPHGVRPKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAFARDLREQDALAV 536

Db 490 VNGHPETKGRAPKAPDAAKP--RVPWFGDKLVADGTKQLLDQLGPKKFAEWENEKRALI 547

QY 537 TDTTTFDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEANGGATYDVAMRFLFEDPDWD 596

Db 548 TDTTMRDGHQSLLATRVRTYDIAPIANAYAQALPNLFSLCEWGGATFDVSMRFLTEDPWE 607

QY 597 RLDELREAMPNVNIQMLLRGNRTVGYTPYDPSVCRAFVKEAASSGVDFIRIFDALNDVSQ 656

Db 608 RLALVREGAPNLLQLMLLRGANGVGYSYDPDNVVKYFVREARAGIDLFRVFDLSLNWVEN 667

QY 657 MRPAIDAVLETNTAVAEVAMAYSGDLSDPNEKLYTLDYLYLKMAEEIVKSGAHILAIKDMA 716

Db 668 MRVMDAVLEEN-KICEAAICYTGDILNPDRAKYDNLNYYVNLAKEVEKAGAHIIAVKDMA 726

QY 717 GLLRPAAVTKLVLTALRRREFDLPVHVHTHTTAGGQLATYFAAAQAGADAVDASAPLSGTT 776

Db 727 GLLKPAARVLFKALREETDLPHFHTHTDTSIGISAATVLAADAGVDVVDAAAMDALSNT 786

QY 777 SQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEIP 836

Db 787 SQPCLGSIIVEALHGSESDGLDPLRIRISFYWEAVRHQYAAAFESDLKGPASEVYLHEMP 846

QY 837 GGQSLNLRQAATAGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAGVD 896

Db 847 GGQFTNLKEQARSLSGLETRWHEVAQAYADVNRMFQDIVKVTTPSSKVVGDMALMMVAQDLT 906

QY 897 PADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSEGGKAPLTEVPEE--E 954

Db 907 VADVNPAKDIAFPDSVVMRMRGDLGQPPSGWPEALQKKVL-----KDEKPFTRVPGSLLP 962

QY 955 QAHLDA-----DDSKERRNS---LNRLLFPKPTEEFLEHRRRRFNGNTSALDDREFFYGLV 1005

Db 963 AADLDAERKSFEDSVGRKLSDOEFASALMYPKVFTDYATAHETYGTSVLPTPVYFYGLK 1022

QY 1006 EGRETILRLPVRTPLLVRLDAISEPDDKMRNVVANVNGQIRPMRVRDRSVESVTATAE 1065

Db 1023 PEEEVFVDLERGKTLVIVN-QAMSETDEKGMVTVPFELNGQPRRIKVPENRAKGASGGVRR 1081

QY 1066 KADSSNKGHVAAAPPAGVV-TVTVAEGDEVKAGDAVAIIEMKMEATITASVDGKIDRVVV 1124

Db 1082 KVEAGNDKQVGAPMPGIVSTVAVVAGQVTKQGDVLLSIEAMKMETAIHAERDGTIAEVLV 1141

QY 1125 PAATKVEGGDLIVV 1138

Db 1142 RPGEQIDAKOLLIV 1155

RESULT 14

T20346

pyruvate carboxylase (EC 6.4.1.1) D2023.2 [similarity] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000

C;Accession: T20346

R;Kershaw, J.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19260

A;Accession: T20346

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1175 <WIL>

A;Cross-references: EMBL:Z81052; PIDN:CAB02872.1; GSPDB:GN00023; CESP:D2023.2

A;Experimental source: clone D2023

C;Genetics:

A;Gene: CESP:D2023.2

A;Map position: 5

A;Introns: 34/3; 103/3; 246/1; 451/3; 567/2; 821/1; 1045/3; 1126/3

C;Superfamily: pyruvate carboxylase; biotin carboxylase homology; lipoyl/biotin-bindin

C;Keywords: ligase

F;1140/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 42.6%; Score 2464; DB 2; Length 1175;

Best Local Similarity 45.2%; Pred. No. 1.4e-127;

Matches 523; Conservative 200; Mismatches 389; Indels 46; Gaps 14;

QY 12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSPASEAVRIGTEGSPVKAY 71

Db 32 FNKVMVANRGEIAIRVFRALTTELNKTSAIYAEQDKNSVHRLKADEAYLVGKGLPPVAAY 91

QY 72 LDIDEIIGAAKVKADAIYPGYGFLSENAQLARECAENGITFCPTPEVLDLTGDKSRVAV 131

Db 92 LTIDQIILETALKHNIDAIHPFGYGLFSERSDFAAACQONAGIVFGPSPDVMARMGDKVAAR 151

Db 321 EEITGIDIVAAQIOIAAGASLPQLGLFQDKITTRGFATQCRITTEDPAKNFQPDGTGRIEV 380

Qy 371 YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVKMTCRGSDFETAVARAQALAEFTVSGVA 429

Db 381 YRSAGNGVRLDGGNAYAGTIIISPHYDSMLVKSCSGSTYIEIVRRKMIRALIEFRIRGVK 440

Qy 430 TNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDYLVN---KP 485

Db 441 TNIPFLLTLLTNPVFIEGTWTTFIDDTLPQFMVSSQNRAQKLLHYLADVAVNGSSIKG 500

Qy 486 HGVPRKDVAAPIDKLPNIKDL-----PLPRGSRDLKQLGPAAFARDLREQDALA 535

Db 501 QIGLPKLSNP--SVPHLHDAQGNVINVTKSAPPSGWRQVLLLEKGAEEFARQVRQFNGTL 558

Qy 536 VTDTTFRDAHQSLLATRVRSFALKPAAEAVAKLTPELLSVEAWGGATYDVAMRFLFEDPW 595

Db 559 LMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGRALECWGGATFDVAMRFLHEDPW 618

Qy 596 DRLDELREAMPNVNIQMLLRGNTVGYTPYPDSVCRAFFVKEAASSGVDFIRIFDALNDVS 655

Db 619 ERLRKLRLSLVPNIPFQMLLRGANGVAYSSLPDINAIDHFVKQAKDNGVDIFRVFDALNDLE 678

Qy 656 QMRPAIDAVLENTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDKM 715

Db 679 QLKVGVDVAVKAG-GVVEATVCFSGDMLQPGKK-YNLDYYLEIAEKIVQMGTILGIKDM 736

Qy 716 AGLLRPAAVTKLVTLRREF-DLPVHVHTHTAGGQLATYFAAAQACADAVDGASAPLSG 774

Db 737 AGTMKPAAAKLLIGSLRAKYPDLPPIHVHTHDSAGTAVASMTACALAGADVVDVAINMSG 796

Qy 775 TTSQPSLSAIVAAFAHTERDGTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHE 834

Db 797 LTSQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMRLLYSCEFADLKGPDPEVIQHE 855

Qy 835 IPGQSLNLRAQATALGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVVGDLALHLVGAG 894

Db 856 IPGQLTNLLFQAQQLGLGEQWAETKRAYREANYLLGDIVKVTPTSXVVGDLAQFMVSNK 915

Qy 895 VDPADFAADPKYDIPDSVIAFLRGELGNPPGHWPEPLRTRALEGRSECKAPLTEVPEEE 954

Db 916 LTSDDVRRRLANSLOFPDSVMDFFEGGLIQPYGGPEPFRSDVLRNK---RRKLTCTRPGL 972

Qy 955 QAHLDADDSKERRNSLNL-----LFPKPTEEFLEHRRRFGNTSALDDREFFY 1002

Db 973 LEPPDLE--KIREDLQNRFGDVDECDVASYNMYPRVYEDFQKMETYGLSVLPTRSFLS 1030

Qy 1003 GLVEGRETLIRLPDVRTPLLVRLDAISEPDDK-GMRNVVANVNGQIRPMRVRDRSVESVT 1061

Db 1031 PLETDEEIEVWIEQGKT-LIIKLQAVGDLNKKTGSEVYFDLNGEMRKIRVADRQKQVET 1089

Qy 1062 ATAEKADSSNKGHVAAPEFAGV-VTVTVAEGDEVKAGDAVAIIIEAMKMEATITASVDGKID 1120

Db 1090 VTKSKADMHDPLHIGAPNAGVIVEVKVHKGLIKKGPVAVLSAMKMEMIISPSDGGQVK 1149

Qy 1121 RVVVPAAATKVEGGDLIVV 1139

Db 1150 EVFVSDGENVDSSDLLVLL 1168

Search completed: March 24, 2004, 22:44:53
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 22:36:41 ; Search time 17 seconds
(without alignments)
3491.763 Million cell updates/sec

Title: US-10-045-072-2
Perfect score: 5788
Sequence: 1 MSTHTSSTLPFAFKKILVANR.....RVVVPAAATKVEGGDLIVVVS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2541.5	43.9	1178	1	PYC_MOUSE
2	2525.5	43.6	1178	1	PYC_RAT
3	2520.5	43.5	1178	1	PYC_HUMAN
4	2457	42.4	1178	1	PYC1_YEAST
5	2436	42.1	1180	1	PYC2_YEAST
6	2391	41.3	1189	1	PYC_PICPA
7	1079	18.6	501	1	PYCA_METJA
8	1036.5	17.9	447	1	ACCC_ANASP
9	1019	17.6	506	1	PYCA_ARCFU
10	997	17.2	491	1	PYCA_HAEIN
11	946	16.3	448	1	ACCC_HAEIN
12	938	16.2	449	1	ACCC_ECOLI
13	936	16.2	449	1	ACCC_ECO57
14	924	16.0	449	1	ACCC_PSEAE
15	911	15.7	725	1	MCCA_MOUSE
16	909	15.7	717	1	MCCA_HUMAN
17	883.5	15.3	654	1	BCCA_MYCTU
18	879	15.2	448	1	ACCC_BACSU
19	875.5	15.1	598	1	BCCA_MYCLE
20	871.5	15.1	567	1	PYCB_METJA
21	866	15.0	703	1	PCCA_HUMAN
22	859.5	14.8	731	1	MCCA_SOYBN
23	859	14.8	734	1	MCCA_ARATH
24	843.5	14.6	704	1	PCCA_RAT
25	835.5	14.4	1835	1	DURL_YEAST
26	825	14.3	568	1	PYCB_METTH
27	804	13.9	590	1	DCOA_SALTY
28	799.5	13.8	595	1	DCOA_KLEPN
29	612.5	10.6	2345	1	COAL_RAT
30	609.5	10.5	2324	1	COAL_CHICK
31	609.5	10.5	2346	1	COAL_HUMAN
32	608.5	10.5	2346	1	COAL_BOVIN
33	601	10.4	2346	1	COAL_SHEEP
1	2541.5	43.9	1178	1	PYC_MOUSE
2	2525.5	43.6	1178	1	PYC_RAT
3	2520.5	43.5	1178	1	PYC_HUMAN
4	2457	42.4	1178	1	PYC1_YEAST
5	2436	42.1	1180	1	PYC2_YEAST
6	2391	41.3	1189	1	PYC_PICPA
7	1079	18.6	501	1	PYCA_METJA
8	1036.5	17.9	447	1	ACCC_ANASP
9	1019	17.6	506	1	PYCA_ARCFU
10	997	17.2	491	1	PYCA_HAEIN
11	946	16.3	448	1	ACCC_HAEIN
12	938	16.2	449	1	ACCC_ECOLI
13	936	16.2	449	1	ACCC_ECO57
14	924	16.0	449	1	ACCC_PSEAE
15	911	15.7	725	1	MCCA_MOUSE
16	909	15.7	717	1	MCCA_HUMAN
17	883.5	15.3	654	1	BCCA_MYCTU
18	879	15.2	448	1	ACCC_BACSU
19	875.5	15.1	598	1	BCCA_MYCLE
20	871.5	15.1	567	1	PYCB_METJA
21	866	15.0	703	1	PCCA_HUMAN
22	859.5	14.8	731	1	MCCA_SOYBN
23	859	14.8	734	1	MCCA_ARATH
24	843.5	14.6	704	1	PCCA_RAT
25	835.5	14.4	1835	1	DURL_YEAST
26	825	14.3	568	1	PYCB_METTH
27	804	13.9	590	1	DCOA_SALTY
28	799.5	13.8	595	1	DCOA_KLEPN
29	612.5	10.6	2345	1	COAL_RAT
30	609.5	10.5	2324	1	COAL_CHICK
31	609.5	10.5	2346	1	COAL_HUMAN
32	608.5	10.5	2346	1	COAL_BOVIN
33	601	10.4	2346	1	COAL_SHEEP

34	561.5	9.7	2483	1	COA2_HUMAN	O00763 homo sapien
35	556.5	9.6	2273	1	HFA1_YEAST	P32874 saccharomyc
36	537.5	9.3	2233	1	COAC_YEAST	Q00955 saccharomyc
37	511	8.8	2280	1	COAC_SCHPO	P78820 schizosacch
38	267.5	4.6	1076	1	CARB_ARCFU	O28994 archaeoglob
39	264	4.6	1060	1	CARB_METTH	O27077 methanobact
40	255	4.4	1072	1	CARB_ECO57	Q8xa38 escherichia
41	255	4.4	1072	1	CARB_ECOLI	P00968 escherichia
42	254.5	4.4	1077	1	CARB_VIBVU	Q8dem2 vibrio vuln
43	253.5	4.4	1099	1	CARB_THEMEA	Q9wz27 thermotoga
44	251.5	4.3	1074	1	CARB_SALTY	P14846 salmonella
45	251.5	4.3	1077	1	CARB_VIBPA	Q87sf3 vibrio para

ALIGNMENTS

RESULT 1
PYC_MOUSE
ID PYC_MOUSE STANDARD; PRT: 1178 AA.
AC Q05320; 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
DE carboxylase) (PCB).
GN PC OR PCX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipocyte;
RX MEDLINE=93189578; PubMed=8446588;
RA Zhang J., Xia W.L., Brew K., Ahmad F.;
RT "Adipose pyruvate carboxylase: amino acid sequence and domain
RT structure deduced from cDNA sequencing";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1766-1770(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lcquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
CC involving the ATP-dependent carboxylation of the covalently
CC attached biotin in the first step and the transfer of the
CC carboxyl group to pyruvate in the second. Catalyzes in a tissue
CC specific manner, the initial reactions of glucose (liver, kidney)
CC and lipid (adipose tissue, liver, brain) synthesis from pyruvate.
CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3) (-) = ADP + phosphate +
CC oxaloacetate.
CC -!- COFACTOR: Biotin and manganese.
CC -!- PATHWAY: Gluconeogenesis and lipogenesis.

Db 973 RIEG-GPGASLPPLNLKEKLDLDRHGEEVTPEDVLSAAMYPDVFAQFKDFTATFGPLD 1031
Qy 994 ALDDREFFYGLVEGRETLIRLPDVRTPLLVRLDAISEPDDKGRNVVANVNGQIRPMVR 1053
Db 1032 SLNTRLFLOQPKIAEEFEVELERGKT-LHIKALAVSDLNRAQORQVFELNGQLRSILVK 1090
Qy 1054 DRVESVTATAEKADSNKNGHVAAPFAG-VVTVTVAEGDEVKAGDAVAIEAMKMEATIT 1112
Db 1091 DTQAMKEMHFHPKALKDKVKGIGAPMPGKVIDVKAAGAKVVKGOPLCVLSAMKMETVVT 1150
Qy 1113 ASVDGKIDRVVVPAAATKVEGDLIVV 1139
Db 1151 SPNEGTRKRVHTKMTLEGDDLILEI 1177

RESULT 3

PYC_HUMAN STANDARD; PRT; 1178 AA.
AC P11498; Q16705;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pyruvate carboxylase, mitochondrial precursor (EC 6.4.1.1) (Pyruvic
carboxylase) (PCB).
GN PC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney, and Liver;
RX MEDLINE=95002202; PubMed=7918683;
RA Wexler I.D., Du Y., Lisgaris M.V., Mandal S.K., Freytag S.O.,
RA Yang B.-S., Liu T.-C., Kwon M., Patel M.S., Kerr D.S.;
RT "Primary amino acid sequence and structure of human pyruvate
carboxylase.";
RL Biochim. Biophys. Acta 1227:46-52(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=94324922; PubMed=8048912;
RA Mackay N., Rigat B., Douglas C., Chen H.S., Robinson B.H.;
RT "CDNA cloning of human kidney pyruvate carboxylase.";
RL Biochem. Biophys. Res. Commun. 202:1009-1014(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney, and Liver;
RA Walker M.E., Jitrapakdee S., Val D.L., Wallace J.C.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 1083-1178 FROM N.A.
RX MEDLINE=87212051; PubMed=3555348;
RA Lamhonwah A.-M., Quan F., Gravel R.A.;
RT "Sequence homology around the biotin-binding site of human propionyl-
CoA carboxylase and pyruvate carboxylase.";
RL Arch. Biochem. Biophys. 254:631-636(1987).
RN [6]
RP SEQUENCE OF 1135-1178 FROM N.A.
RX MEDLINE=85030380; PubMed=6548474;
RA Freytag S.O., Collier K.J.;
RT "Molecular cloning of a cDNA for human pyruvate carboxylase.
Structural relationship to other biotin-containing carboxylases and
regulation of mRNA content in differentiating preadipocytes.";
RL J. Biol. Chem. 259:12831-12837(1984).
RN [7]
RP VARIANTS PC DEFICIENCY THR-610 AND ILE-743.
RX MEDLINE=98254451; PubMed=9585612;
RA Carbone M.A., MacKay N., Ling M., Cole D.E.C., Douglas C., Rigat B.,
RA Feigenbaum A., Clarke J.T.R., Haworth J.C., Greenberg C.R.,
RA Seargeant L., Robinson B.H.;
RT "Amerindian pyruvate carboxylase deficiency is associated with two
distinct missense mutations.";
RL Am. J. Hum. Genet. 62:1312-1319(1998).
RN [8]
RP VARIANTS PC DEFICIENCY ALA-145 AND CYS-451.
RX MEDLINE=98244401; PubMed=9585002;
RA Wexler I.D., Kerr D.S., Du Y., Kaung M.M., Stephenson W., Lusk M.M.,
RA Wappner R.S., Higgins J.J.;
RT "Molecular characterization of pyruvate carboxylase deficiency in two
consanguineous families.";
RL Pediatr. Res. 43:579-584(1998).
CC -!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
involving the ATP-dependent carboxylation of the covalently
attached biotin in the first step and the transfer of the
carboxyl group to pyruvate in the second. Catalyzes in a tissue
specific manner, the initial reactions of glucose (liver, kidney)
and lipid (adipose tissue, liver, brain) synthesis from pyruvate.
CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
oxaloacetate.
CC -!- COFACTOR: Biotin and manganese.
CC -!- PATHWAY: Gluconeogenesis and lipogenesis.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- DISEASE: Defects in PC are the cause of pyruvate carboxylase to
deficiency (PC deficiency) [MIM:266150]. PC deficiency leads to
lactic acidosis, mental retardation and death. It occurs in three
forms: mild or type A, severe neonatal or type B, and a very mild
lacticacidemia.
CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
AND CARBAMYL PHOSPHATE SYNTHETASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; U04641; AAA99537.1; -
DR EMBL; S72370; AAB31500.1; -
DR EMBL; U30891; AAA82937.1; -
DR EMBL; BC011617; AAH11617.1; -
DR EMBL; M26122; AAA36423.1; -
DR EMBL; K02282; AAA60033.1; -
DR PIR; G01933; JC2460.
DR HSSP; P24182; 1BNC.
DR Genew; HGNC:8636; PC.
DR GK; P11498; -
DR MIM; 266150; -

DR GO:0005524; F:ATP binding; TAS.
DR GO:0009374; F:biotin binding; TAS.
DR GO:0004736; F:pyruvate carboxylase activity; TAS.
DR InterPro: IPR001882; Biotin BS.
DR InterPro: IPR005482; Biotin_carb_C.
DR InterPro: IPR000089; Biotin_lipoYL.
DR InterPro: IPR005479; CPase_L_D2.
DR InterPro: IPR005481; CPase_L_N.
DR InterPro: IPR000891; HMGL-like.
DR InterPro: IPR003379; PYC_OADA.
DR InterPro: IPR005930; Pyruv carbox.
DR Pfam: PF02785; Biotin carb_C; 1.
DR Pfam: PF00364; biotin_lipoYL; 1.
DR Pfam: PF00289; CPSase_L chain; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR Pfam: PF00682; HMGL-like; 1.
DR Pfam: PF02436; PYC_OADA; 1.
DR TIGRFAMs: TIGR01235; pyruv carbox; 1.
DR PROSITE: PS00188; BIOTIN; 1.
KW Ligase; Multifunctional enzyme; Biotin; Manganese; Gluconeogenesis;
KW ATP-binding; Mitochondrion; Lipid synthesis; Transit peptide;
KW Disease mutation.
FT TRANSIT 1 20 MITOCHONDRION (POTENTIAL).
FT CHAIN 21 1178 PYRUVATE CARBOXYLASE.
FT DOMAIN 21 549 BIOTIN CARBOXYLASE (BY SIMILARITY).
FT DOMAIN 550 1000 CARBOXYLTRANSFERASE (BY SIMILARITY).
FT DOMAIN 1096 1178 BIOTIN CARBOXYL CARRIER PROTEIN (BY SIMILARITY).
FT NP BIND 198 203 ATP (BY SIMILARITY).
FT ACT_SITE 328 328 BY SIMILARITY.
FT BINDING 1144 1144 BIOTIN (BY SIMILARITY).
FT VARIANT 145 145 V -> A (in PC deficiency; mild).
FT VARIANT 451 451 /FTID=VAR_015199.
FT VARIANT 451 451 R -> C (in PC deficiency; mild).
FT VARIANT 610 610 /FTID=VAR_015200.
FT VARIANT 610 610 A -> T (in PC deficiency; mild).
FT VARIANT 743 743 /FTID=VAR_008095.
FT VARIANT 743 743 M -> I (in PC deficiency; mild).
FT CONFLICT 225 226 /FTID=VAR_008096.
FT CONFLICT 352 352 LA -> WP (IN REF. 2).
FT CONFLICT 385 386 A -> S (IN REF. 3).
FT CONFLICT 486 487 RS -> PT (IN REF. 2).
FT CONFLICT 638 638 EL -> DV (IN REF. 2).
FT CONFLICT 729 729 P -> R (IN REF. 2).
FT CONFLICT 774 775 E -> A (IN REF. 2).
FT CONFLICT 774 775 DT -> AP (IN REF. 2).
SQ SEQUENCE 1178 AA; 129633 MW; 381F527553A20095 CRC64;

Query Match 43.5%; Score 2520.5; DB 1; Length 1178;
Best Local Similarity 46.3%; Pred. No. 3.4e-127;
Matches 537; Conservative 183; Mismatches 412; Indels 29; Gaps 12;

QY 6 SSTLPA-----PKKILVANRGEIAVRAAPRAALETGAATVAIYPREDRGSFHR 53
Db 19 TSTAPASPNVRRLEYKPIKKVMVANRGEIAIRVRACTELGIRTVAIYSEQDTGQMRHQ 78

QY 54 FASEAVRIGTEGSPVKAYLDIDEIIGAAKVKADAIYPGYGFLSENAQLARECAENGITF 113
Db 79 KADEAYLIGRLAPVQAYLHIPDIIKVAKENNVDAVHPGYGFLSERADFAQACQDAGVRF 138

QY 114 IGPTPEVLDLTGDKSRVTAACKAGLPVL-AESTPSKNIDEIVKSAEGQTYPIFVKAVAG 172
Db 139 IGPSPEVVRKMGDKVEARAIAAAGVPVPGTDAPITSLHEAHEFSNTYGPPIIFKAAAG 198

QY 173 GCGRGVRFVASPELRLKLAATEASREAEAAFGDGAUVVERAVINPQHIEVQILGDHTGEVV 232
Db 199 GCGRGVRFVHSYEELEENYTRAYSEALAAFGNGALFVEKFKPRHIEVQILGDQYGNIL 258

QY 233 HLYERDCSLORRHQKVEIAPACHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKG 292
Db 259 HLYERDCSIORRHQKVEIAPAAHLDPLRLTRLTSDSVKLAQGVYENAGTVEFLVDRHG 318

QY 293 NHVFIEMNPRIQVHTVTEVTEVDLVKAQMRLLAAGATLKELGLTQDKIKTHGAALQCRI 352

Db 319 KHYFIEVNSRLQVEHTVTEETITDVLVHAQIHVAEGRSLPDLGRQENIRINGCAIQCRV 378
QY 353 TTEDPNNGFRPDTGTITAYRSPGGAGVRLDGAQAQL-GGEITAHFDSMLVKMTCRGSDFET 411
Db 379 TTEDPARSFQDPTGRIEVRSGEGMIRLDNASAFQGAIVISPHYDSLIVKVIAHKGDHPT 438
QY 412 AVARAQRALAEFTVSGVATNIGFLRALLREEDFTSKRIATGFTADHPHLLQAPPADDEQG 471
Db 439 AATKMSRALAEFRVGVKTNIAFLQNVLNNQQFLAGTVDTQFIDENPELFQLRPAQNRAQ 498
QY 472 RILDYLAADVTVNKPFGVRP-KDVAAPIDK-LPNIKDLPLPRGSRDLKQLGPAAFARDLR 529
Db 499 KLLHYLGHVMVNGPTTPIPVKASPSPTDPVPAVPIGPPAGPFRDILLREGPEGFARAVR 558
QY 530 EQDALAVTDTTFRDAHQSLLATRVRSFALKPAAEAVALTPELLVSVEAWGGATYDVAMRF 589
Db 559 NHPGLLLMDTTFRDAHQSLLATRVTRTHDLKKIAPYVAHNFSKLFSMENWGGATFDVAMRF 618
QY 590 LFEDPWDRDLDELREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFKVEKAASSGVDFIRIFD 649
Db 619 LYECPWRRRLQELRELIPNIPFQMLLRGANAVGYTNPYDNVVFKECEVAKENGMDFRVFD 678
QY 650 ALNDVSQMRPAIDAVLENTNTAFAEVNMAVSGDSDPNKLYTLDYILKMAEEIVKSGAHI 709
Db 679 SLNYLPNMLLGMEEAGSAG-GVVEAAISYTGADVADPSRTKYSLOYIMGLABELVRAGTHI 737
QY 710 LAIKDMAGLLRPAAVTKLVTAALREF-DLPVHVHTHTAGGQLATYFAAAQAGADAVDGA 768
Db 738 LCIKDMAGLLKPTACTMLVSSLRDRFPDPLPHIHTHTDSGAGVAAMLACAQAGADVVDVA 797
QY 769 SAPLSGTTSQPSLSAIVAAFAHAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGTG 828
Db 798 ADMSGMTSQPSMGALVACTRGTPLDTEVPMERVFYSEYEWEGARGLYAADFCTATMKSG 857
QY 829 R--VYRHEIPGGQLSNLRAQATAGLADRFELIEDNYAAVNEMLGRPTKVTPSSKVVGDL 886
Db 858 NSDVYENEIPGGQYTNLHFQAHSMGLSKFKEVKKAYVEANQMLGDLIKVTPSSKIVGDL 917
QY 887 ALHLVGAGVDPADFAADPQKYDIPDSVIAFLRGELGNPPGGWPEPLRTRAL-----EGR 940
Db 918 AQFMVQNGLSRAEBAEAAEELSFPFRSVVEFLQGYIGVPHGGFPEPPFRSKVLKDLPRVEGR 977
QY 941 SEGKAPLTVPEPEEQAHLDA-DDSKERRNSLNRLLPKPTEEFLEHRRRFRGNTSALDDRE 999
Db 978 PGASLPPLDLQALEKELVDRHGEVTPEDVLSAAMYPDVFAHFKDFTATFGPLDSLNTRL 1037
QY 1000 FFYGLVEGRETLIRLPDVRTPLLVRLDAISEDDDKMRNVVANVNGQIRPMRVDRSVES 1059
Db 1038 FLQGPKEAEFEFELEERKXT-LHIKALAVSDLNRAQORQVFFELNGQLRSILVKDTQAMK 1096
QY 1060 VTATAEKADSSNKGHVAAFPAG-VVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDGK 1118
Db 1097 BMHFHPKALKDVKGQIGAPMPGKVIDIKVAVAGAKVAKGQPLCVLSAMKMETVVTSPMEGT 1156
QY 1119 IDRVVVPAATKVEGGDLIVVV 1139
Db 1157 VRKVHVTKDMTLEGDDLILEI 1177

RESULT 4
PYC1 YEAST
ID PYC1 YEAST STANDARD; PRT; 1178 AA.
AC P11154;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pyruvate carboxylase 1 (EC 6.4.1.1) (Pyruvic carboxylase 1) (PCB 1).
GN PYC1 OR PYV OR YGL062W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae;
OC NCBI_TaxID=4932;

[1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=88298805; PubMed=3042770;
RA Lim F., Morris C.P., Occhiodoro F., Wallace J.C.;
RT "Sequence and domain structure of yeast pyruvate carboxylase.";
RL J. Biol. Chem. 263:11493-11497(1988).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C;
RX MEDLINE=97377993; PubMed=9234674;
RA Feuermann M., de Montigny J., Potier S., Souciet J.-L.;
RT "The characterization of two new clusters of duplicated genes
RT suggests a 'lego' organization of the yeast Saccharomyces cerevisiae
RT chromosomes.";
RL Yeast 13:861-869(1997).
[3]
RN SEQUENCE OF 1003-1178 FROM N.A.
RX MEDLINE=87241529; PubMed=3036126;
RA Morris C.P., Lim F., Wallace J.C.;
RT "Yeast pyruvate carboxylase: gene isolation.";
RL Biochem. Biophys. Res. Commun. 145:390-396(1987).
CC -!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
CC involving the ATP-dependent carboxylation of the covalently
CC attached biotin in the first step and the transfer of the
CC carboxyl group to pyruvate in the second.
CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3) (-) = ADP + phosphate +
CC oxaloacetate.
CC -!- COFACTOR: Biotin and zinc.
CC -!- PATHWAY: Gluconeogenesis.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC AND CARBAMYL PHOSPHATE SYNTHETASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; J03889; AAA34843.1; -.
DR EMBL; Z72584; CAA96765.1; -.
DR PIR; S64066; QYBYP.
DR HSSP; P24182; 1BNC.
DR GerMOnline; 141110; -.
DR SGD; S0003030; PYC1.
DR GO; GO:0005829; C:cytosol; IDA.
DR GO; GO:0004736; F:pyruvate carboxylase activity; IDA.
DR InterPro; IPR001882; Biotin_BS.
DR InterPro; IPR005482; Biotin_carb_C.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR000891; HMGL-like.
DR InterPro; IPR003379; PYC_OADA.
DR InterPro; IPR005930; Pyruv_carbox.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF00289; CPase_L_chain; 1.
DR Pfam; PF02786; CPase_L_D2; 1.
DR Pfam; PF00682; HMGL-like; 1.
DR Pfam; PF02436; PYC_OADA; 1.
DR TIGRFAMs; TIGR01235; pyruv_carbox; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00866; CPASE_1; 1.
DR PROSITE; PS00867; CPASE_2; 1.
KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
KW Zinc; Multigene family.
FT NP_BIND 182 187 ATP (POTENTIAL).
FT ACT_SITE 312 312 BY SIMILARITY.
FT BINDING 1135 1135 BIOTIN (BY SIMILARITY).

FT	CONFLICT	462	462	T -> G (IN REF. 1).	
FT	CONFLICT	493	493	V -> D (IN REF. 1).	
FT	CONFLICT	595	595	R -> A (IN REF. 1).	
FT	CONFLICT	619	619	E -> Q (IN REF. 1).	
FT	CONFLICT	664	664	G -> S (IN REF. 1).	
FT	CONFLICT	772	772	A -> R (IN REF. 1).	
FT	CONFLICT	879	879	E -> Q (IN REF. 1).	
FT	CONFLICT	909	909	Q -> K (IN REF. 1).	
SQ	SEQUENCE	1178	AA; 130099	MW; BC7110A8AFB23E04	CRC64;

Query Match 42.4%; Score 2457; DB 1; Length 1178;
Best Local Similarity 46.7%; Pred. No. 8.5e-124;
Matches 541; Conservative 173; Mismatches 401; Indels 44; Gaps 16;

QY	14	KILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRSPASEAVRIGTEG--SPVKAY	71
Db	21	KILVANRGEIPIRIFRTAHELSTHSHEDRLSTHKQKADAEYVIGEVGYTTPVGAY	80
QY	72	LDIDEIIGAACKVKADAIYPGCGFLSENAQLARECAENGITFPGTPEVLDLTGDKSRV	131
Db	81	LAIDEIISIAQKHQVDFIHPGYGELSENSEFADKVVKAGITWIGPPAEVIDSVGDKVSAR	140
QY	132	TAACKAGLPVLAEST-PSKNIDEIVKSAEGQTYPIFVKAVAGGGGRGMRFVSPDELRLK	190
Db	141	NLAAKANVETVPGTPGPIETVEEALDFVNEYGYPIIKAFGGGGGRGMRVREGDDVADA	200
QY	191	ATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVVHLYERDCSLQRHQQVVE	250
Db	201	FORATSEARTAFNGTCTCFERFLDKPKHIEVQLLADNHNHGVVHLFERDCSVQRRHQQVVE	260
QY	251	IAPAQHLDPELRDRICADAVKFCRSIGYQGAGTVEFLVDEKGNHVFIEPNRIQVEHTVT	310
Db	261	VAPAKTLPREVRDAILTDVAKLAKECGYRNAGTAFLVDNQRHRYFIEINPRIQVEHTIT	320
QY	311	EEVTEVDLVKAQMRLLAAGATLKLGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTITA	370
Db	321	EEITGIDIVAQAQIAAGASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDGTGRIEV	380
QY	371	YRSPGGAGVRLDGA-AQLGGEITAHFDSMLVMTCRGSDFFETAVARAQALAEFTVSGVA	429
Db	381	YRSAGNGVRLDGGNAYAGTIIISPHYDSMLVKCSCSGSTYEIVRRKMIRALIEFIRGVK	440
QY	430	TNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADDEQGRILDYLDVTVN---KP	485
Db	441	TNIPFLTLITNPVFIETGYTWTTFIDDTPLQFMVSSQNRAQKLHLHYLADVAVNGSSIKG	500
QY	486	HGVRPKDVAAPIDKLPNIKOL-----PLPRGSRDLKQLGPAAAFARDLREQDALA	535
Db	501	QIGLPKLKSNP--SVPHLHDAQGNVINVTKSAPPSGWRQVLEKGPFAEFARQVRQFNGTL	558
QY	536	VTDTTFRDAHQSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGATYDVAMRFLFEDPW	595
Db	559	LMDTTWRDAHQSLLATRVTRTHDLATTAPTTAHALAGRALECGCATFDVAMRFLHEDPW	618
QY	596	DRLDELREAMPNVNIQMLLRGNTVGYTPYDPSVCRAFVKEAASSGVDFIRIFDALNDVS	655
Db	619	ERLRKLSLVNPVNIQMLLRGANGVAYSSLPDNDIDHFVKQAKDNGVDIFRVFDALNDLE	678
QY	656	QMRPAIDAVLENTAVAEVAMAYSGDLSDPNEKLYTLDYLLKMAEEIVKSGAHILAIDKM	715
Db	679	QLKVGVDVAVKKAG-GVVEATVCFSGDMLQPKKK-YNLDYYLEIAEKIVQMGMTHILGIKDM	736
QY	716	AGLLRPAAVTKLVATLRRF-DLPVHVHTHTAGGQLATYFAAAAGADAVDGASAPLSG	774
Db	737	AGTMKPAAAKLLIGSLRAKYPDLPPIHVHTHDSAGTAVASMTACALAGADVVDVAINMSG	796
QY	775	TTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPEYWEAVRGLYLPPFESGTPGTRVYRHE	834
Db	797	LTSQPSINALLASL-EGNIDTGINVHVRELDAYWAEMLLYSCFEADLKGPDPVEVQHE	855
QY	835	IPGQLSNLRACATAGLADRFELIEDNYAAVNEMLGRPTKVTSSKVVGDALHLVAG	894
Db	856	IPGQLTNLLFQAQQLGLGEQWAEETKRAYREANYLLGDIVKVTFTSKVVGDLAQFMVSNK	915

QY 244 RHQKVEIAPACHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEINPRI 303
DB 255 RHQKVEVAPAKTLPREVRDAILTDVAKLAKVGYRNAGTABFLVDNQNRHYFIEINPRI 314
QY 304 QVEHTVTEEVTEVDLVKAQMRLLAAGATLKLGLTQDKIKTHGAALQCRITTEDFNNGFRP 363
DB 315 QVEHTITEETIGIDIVSAQIQIAGATLTQLGLLQDKITTRGFSIQCRITTEDPSKNFQP 374
QY 364 DTGTITAYRSPGGAGVRLDGA-AQLGGEITAHFDSMLVXMTCRGSDFFETAVARACRALAE 422
DB 375 DTGRLEVRSAGNGVRLDGGNAYAGATISPHYDSMLVKSCSGSTYIEIVRRKMRALIE 434
QY 423 FTVSGVATNIGFLRALLREEDFTSKRIATGFIADHPHLLQAPPADDEQGRILDYLDVTV 482
DB 435 FRIRGVKTNIPFLLTLLTNPNVPIEGTYWTFIDDTLPQLFQMVSSQNRAQKLLHYLDLAV 494
QY 483 N----KPHGVPRKDVAAPIDKLPNIKDL-----PLPRGSRDRLLKQLGPAFAFARDL 528
DB 495 NGSSIKGQIGLPLKXSNP--SVPHLHDAQNVINVTKSAPPSGWRQVLLLEKGPSEFAKQV 552
QY 529 REQDALAVTDTTFRDAHQSLLATRVRSFALKPAEAEAVAKLTPELLSVEAWGGATVDVAMR 588
DB 553 RQFNGTLLMDTTWRDAHQSLLATRVTHDLATIAPTTAHALAGAFALCWCWGATFVAMR 612
QY 589 FLFEDPWRDLDELREAMPNVNIQMLLRGNVTGYTPYDPSVCRAPVKEAASSGVDIRIF 648
DB 613 FLHEDPWERLRLKRLSLVPNIPFQMLRGANGVAYSSLPDNDHDFVKQAKONGVDIFRVF 672
QY 649 DALNDVSQVRPAIDAVLENTTAVAEVAMAYSGDLSDPNEKLYLTDYLLQMAEEIVKSGAH 708
DB 673 DALNDLEQLKGVNAVKKAG-GVVEATVCYSGDMLQPGKK-YNLDYLYLEVVEKIVQMGTH 730
QY 709 ILAIKOMAGLLRPAAVTKLVTLALREF-DLPVHVHTHTDAGQLATYFAAAQAGADAVDG 767
DB 731 ILGIKOMAGTKMPAAAKLLIGSLRTRYPDLPIHVSHSDSAGTAVASMTACALAGADVVDV 790
QY 768 ASAPLSGTTQSPLSAIIVAFAHTRDRTGLSLEAVSDLEPWEAVRGLYLPFESGTPGPT 827
DB 791 AINSMSGTLTQPSINALLASL-EGNIDTGINVEHVRELDAYWAEMLRLLYSCFEADLKGPD 849
QY 828 GRVYRHEIPGQSLNLRQAATATGLADRFELIEDNYAAVNMELGRPTKVTTPSSKVVGDLA 887
DB 850 PEVYQHEIPGQSLNLLFQAQQLGLGEQWAEKTRAYREANYLLGDIVKVTPTSKVVGDLA 909
QY 888 LHLVGAGVDPADFAADPKYDIPDSVIAFLRGLGNPPGWPPEPLRTRALEGRSEKAPL 947
DB 910 QFMVSNKLTSDDIRRLANSDFPDSYNDFFEGILQYQYGGFPEPLRSDVLRNK---RRKL 966
QY 948 TEVPEEEQALHDADDKSKERRNSLNL-----LFPKPTTEEFLEHRRRRFGNTSAL 995
DB 967 TCRPGLLEPEPDL--KIREDLQNRFGDIDECDVASYNMYPVYEDFQKIRETYGDLVL 1024
QY 996 DDREFFYGLVEGRETILRLPDVRTPLVRLDAISEPDDK-GMRNVANVNGQIRPMVRD 1054
DB 1025 PTKNFLAPAEPEDEIEVTIEQGT-LIIKLQAVGDLNKKTGQREYVFELNGELRKLIRVAD 1083
QY 1055 RSVESVTATAEKADSSNKGHVAAPFAGV-VTVTVAEAGDEVKAGDAVAIIIEAMKMEATITA 1113
DB 1084 KSONIQSVAKPKADVHDTHQIGAPMAGVIEVKVHKGSLVKKGESIAVLSAMKMEMVVS 1143
QY 1114 SVDGKIDRVVVPAAATKVEGGDLIVVV 1139
DB 1144 PADGQVKDVFIKDGESVDASDLLVVL 1169

RESULT 6
ID PYC PICPA STANDARD; PRT; 1189 AA.
AC P78992;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Pyruvate carboxylase (EC 6.4.1.1) (Pyruvic carboxylase) (PCB).

GN PYC1.
OS Pichia pastoris (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4922;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98301182; PubMed=9639311;
RA Menendez J., Delgado J., Gancedo C.;
RT "Isolation of the Pichia pastoris PYC1 gene encoding pyruvate
carboxylase and identification of a suppressor of the pyc
phenotype.";
RL Yeast 14:647-654(1998).
CC -1- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
involving the ATP-dependent carboxylation of the covalently
attached biotin in the first step and the transfer of the
carboxyl group to pyruvate in the second (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3)(-) = ADP + phosphate +
oxaloacetate.
CC -1- COFACTOR: Biotin and zinc.
CC -1- PATHWAY: Gluconeogenesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
AND CARBAMYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y11106; CAA71993.1; -.
CC HSSP; P24182; 1DV1.
CC InterPro; IPR001882; Biotin_BS.
CC InterPro; IPR005482; Biotin_carb_C.
CC InterPro; IPR000089; Biotin_lipoyl.
CC InterPro; IPR005483; CPase_L.
CC InterPro; IPR005479; CPase_L_D2.
CC InterPro; IPR005481; CPase_L_N.
CC InterPro; IPR000891; HMGL-Like.
CC InterPro; IPR003379; PYC_OADA.
CC InterPro; IPR005930; Pyruv_carbox.
CC Pfam; PF02785; Biotin_carb_C; 1.
CC Pfam; PF00364; biotin_lipoyl; 1.
CC Pfam; PF00289; CPase_L_chain; 1.
CC Pfam; PF02786; CPase_L_D2; 1.
CC Pfam; PF00682; HMGL-Like; 1.
CC Pfam; PF02436; PYC_OADA; 1.
CC PRINTS; PR00098; CPSASE.
CC TIGRFAMs; TIGR01235; pyruv_carbox; 1.
CC PROSITE; PS00188; BIOTIN; 1.
CC PROSITE; PS00866; CPSASE_1; 1.
CC PROSITE; PS00867; CPSASE_2; 1.
KW Ligase; Multifunctional enzyme; Biotin; Gluconeogenesis; ATP-binding;
Zinc.
FT NP_BIND 185 190 ATP (POTENTIAL).
FT ACT_SITE 315 315 BY SIMILARITY.
FT BINDING 1140 1140 BIOTIN (BY SIMILARITY).
SQ SEQUENCE 1189 AA; 131400 MW; 8B6E858079657914 CRC64;

Query Match 41.3%; Score 2391; DB 1; Length 1189;
Best Local Similarity 45.3%; Pred.No. 2.9e-120;
Matches 526; Conservative 177; Mismatches 423; Indels 34; Gaps 16;

QY 6 SSTLPAPFKKILVANRGEIARAPRAALETGAATVAIYPREDRGSFHRSEAVRIGTEG 65
DB 16 SLLGTWNKILVANRGEIPIRIFRTAHELMSMNTVAIYSHEDRLSMHRLKADEAYVIGERG 75
QY 66 --SPVKAYLDIDEIIGAAGKVKADAIYPGYGFLSENAQLARECAENGITFIGPTPEVLDL 123
DB 76 QYSPVQAYLAIDEIIKIAVKHNVNMTHPGYGCSENSEFAKVEENGILWVGPSDTVIDA 135

QY 124 TGDKSRVATAAKKAGLPVLAEST-PSKNIDEIVKSAEQTYPIFYKAVAGGGGRGMRFFVA 182
Db 136 VGDKVSARNLAYAANVPTVPGPIEDVAQATAFVEEYGYPIVIAAFAAGGGGRGMRVVR 195
QY 183 SPDELRLKLAATESAEAAFGDGVYVERAVINPQHIEVQILGHTGEVVLHYERDCSLQ 242
Db 196 EGDIEDAFLRASBAKTAFGNGTVFIERFLDKPKHIEVQLADNYGNVHLFERDCSVQ 255
QY 243 RRHQKVEIAPAQHLDPELRDRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEMNPR 302
Db 256 RRHQKVARNCSAKTLPEVVRNAILNDVAKLTANYRNAGTAEFLVDSQNRHYFIEINPR 315
QY 303 IQVEHTVTEEVTEVDLVKAQMRLAAGATLKELGLTQDKIKTHGAALQCRITTEDPNNGFR 362
Db 316 IQVEHTITEEITGVDIVAAQIQIAAGASLEQLGLLQEKITTRGFALQCRITTEDPTKNFQ 375
QY 363 PDTGTTITAYRSPGGAGVRLDGAALGGE-ITAHFDSMLVKMTCRGSDPFETAVARAQALA 421
Db 376 PDTGKIEVYRSSGGVRLDGGNGFAGAVISPHYDSMLVKCSTSGSNYEIRRRKMIRALV 435
QY 422 EFTVSGVATNIGFLRALLREEDFTSKRIATGFIADPHLLQAPPADBEQGRILDYLDVLT 481
Db 436 EFRIRGVKTNIPFLALLTHPVFMTSECWTTTFIDDTPELFKILTSQNAQKLLAYLGDLA 495
QY 482 VN-----KPHGVRPKDVAAPIDKLPNKD--LPLPRGSRDRILKQLGPAAFARDL 528
Db 496 VNGSSIKQIGLPLKH--KEADIPSIIDINGVIDVISIPPPDGRWQFLLEKGPQFAQQV 553
QY 529 REQDALAVTDTTFRDAHOSLLATRVRSFALKPAEAAVAKLTPELLSVEAWGGATYDVAMR 588
Db 554 RAPPGLMIMDTTWRDAHQSLLATRVTRTHDLLNTPATSYALHHAFALECGWGFEDVSMR 613
QY 589 FLFEDPWDRDLDELREAMPVNIQMLLRGNRTVGYTPYDPSVCRAFYKVEAASSGVDFIRIF 648
Db 614 FLHEDPWQRLKRLKAVENIPFPMLLRGGNGVAYSLPDNAIDHFLKQAKDITGVDVFRVF 673
QY 649 DALNDVSOMRPAIDAVLENTTAVAEVAMAYSGDLSDPNEKLYTLDDYILKMAEEIVKSGAH 708
Db 674 DALNDIEQLKVGVDVAVKAG-GVVEATMVCYSGDMLKPKGX-YNLEYVINLATEIVEMGTH 731
QY 709 ILAIKDMAGLLRPAAVTKLVLTALRREF-DLPVHVHTHTDTAGQLATYFAAAQAGADAVDG 767
Db 732 ILAVKDMAGTLKPTAAKQJISALRRKFPSPLPPIHVHTHDSAGTVASMVACARAGADVTV 791
QY 768 ASAPLSGTTSQPSLSAIVAAFAHTRRDTGLSLEAVSDLEPYWEAVRGLYLPFESGTPGPT 827
Db 792 RVNMSGWTSQPSMSAFIASL-DGEIETGIPEANAREIDAYWAEMLLSYCFEADLKQPD 850
QY 828 GRVVRHEIPGGQLSNLRAQATLGLADRFELIEDNYAAVNEMLGRPTKVTTPSSKVUGDLA 887
Db 851 PEVYQHEIPGGQLTNLLFQAQVGLGKQVETKKAVEAANRLLGDIVKVTPTSKVUGDLA 910
QY 888 LHLVAGVDPADFAADPQKVDIPDSVIAFLRGELGNPPGGWPEPLRTRALEGRSE---GK 944
Db 911 QFMVSNKLSSEEDVERLASELDFPDSVLDFFEGMLMGTGYGFPPEPLRTNVISGKRRLTSR 970
QY 945 APLTEVPEEEQA---HLDADDSKERRNSL-NRLLFPKPTPEEFLEHRRRFRGNTSALDDREF 1000
Db 971 PGLTLEPNIPAIREDLEAREFSKVTENDVASYNMYPKVYEAAYKKQQLYGLDLSVLPTRNF 1030
QY 1001 FY--GLVVEGRETLIRLPDVRTPLLVRLDAISE-PDDKGNRVNVANVNGQIRPMRVDRSV 1057
Db 1031 LSPPKIDEERHVTIVTITRKTLLIKMAEGELSQSSTREVFYFELNGEMRKVTVEDXNG 1090
QY 1058 ESVTAEAKADSSNKGHVAAFPAGVVT-VTVAEGDEVKAGDAVAIIEMKMEATITASVD 1116
Db 1091 AVETITRPKADAHNPNEIGAPMAGVVVEVRVHENGVEVKGDPIAVLSAMKMEMVISPVA 1150
QY 1117 GKIDRVVVPAAATKVEGGDLI 1136
Db 1151 GRIGQIAVKENDSVSDASDLI 1170

RESULT 7
PYCA_METJA
ID PYCA_METJA STANDARD; PRT; 501 AA.
AC Q58626;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyruvate carboxylase subunit A (EC 6.4.1.1) (Pyruvic carboxylase A).
GN PYCA OR MJ1229.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2561 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
RN [2]
RP SEQUENCE OF 1-12, AND FUNCTION.
RX MEDLINE=21034791; PubMed=11195096;
RA Mukhopadhyay B., Patel V.J., Wolfe R.S.;
RT "A stable archaeal pyruvate carboxylase from the hyperthermophile
Methanococcus jannaschii.";
RL Arch. Microbiol. 174:406-414(2000).
CC -!- FUNCTION: Pyruvate carboxylase catalyzes a 2-step reaction,
CC involving the ATP-dependent carboxylation of the covalently
CC attached biotin in the first step and the transfer of the carboxyl
CC group to pyruvate in the second.
CC -!- CATALYTIC ACTIVITY: ATP + pyruvate + HCO(3) (-) = ADP + phosphate +
CC oxaloacetate.
CC -!- Cofactor: ATP, magnesium (or manganese or cobalt), pyruvate and
CC bicarbonate.
CC -!- ENZYME REGULATION: INHIBITED BY MAGNESIUM, WHEN ITS CONCENTRATION
CC EXCEEDED THE ATP ONE, AND BY HIGH CONCENTRATION OF ATP AND ALPHA-
CC KETOGLUTARATE.
CC -!- PATHWAY: Gluconeogenesis.
CC -!- SUBUNIT: Heterooctamer of four A and four B subunits.
CC -!- MASS SPECTROMETRY: MW=55500; METHOD=WALDI.
CC -!- MISCELLANEOUS: ITS OPTIMUM PH IS 8.5 AND THE OPTIMUM TEMPERATURE
CC IS 80-90 DEGREES CELSIUS.
CC -!- SIMILARITY: WITH OTHER BIOTIN CARBOXYLASES, LIPOAMIDE TRANSFERASES
CC AND CARBAMYL PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67563; AAB99232.1; --
CC PIR; D64453; D64453.
CC HSSP; P24182; 1BNC.
CC TIGR; MJ1229; --
CC InterPro; IPR004549; ACCC.
CC InterPro; IPR005482; Biotin carb_C.
CC InterPro; IPR005483; CPase_L.
CC InterPro; IPR005479; CPase_L D2.
CC InterPro; IPR005481; CPase_L N.
CC Pfam; PF02785; Biotin_carb_C; 1.

```
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR PRINTS; PR00098; CPSASE.
DR TIGRFAMs; TIGR00514; accC; 1.
DR PROSITE; PS00866; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
KW Ligase; Multifunctional enzyme; Gluconogenesis; Magnesium; Pyruvate;
KW ATP-binding; Complete proteome.
FT NP_BIND 162 167 ATP (POTENTIAL).
FT ACT_SITE 291 291 POTENTIAL.
SQ SEQUENCE 501 AA; 55402 MW; 04D2E401892F872F CRC64;

Query Match 18.6%; Score 1079; DB 1; Length 501;
Best Local Similarity 48.2%; Pred. No. 1.3e-50;
Matches 218; Conservative 82; Mismatches 146; Indels 6; Gaps 4;

QY 12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASAVRIGTEGSPVKAY 71
Db 12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASAVRIGTEGSPVKAY 71
QY 72 LDIDEIIGAAGKVKADAIYPGYGFLSENAQLARECAENGITFTGPTPEVLDLTGDKSRV 131
Db 61 LNIIDAILNVAEKAKVDIAHPGYGFLAENAFARAVKAGFEFGPNDAIEMGSKINAK 120
QY 132 TAAKAGLPVLAESTPS-KNIDEIVKSAEQTYPIFVKAVAGGGGRGMRFVASPDRLKL 190
Db 121 KIMKAGVPLIPGSEGAIEDIDEAIEIAEAGFVVKASAGGGGGMVSAVSKELKEV 180
QY 191 ATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVTVHLYERDCSLQRHOKVVE 250
Db 181 IESARNIAKSAFGDPTVIEKYLENPRHIEIQLLGDKHNIHLGDRECSIQRRHOKLIE 240
QY 251 IAPAQHLDPEDLRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEHNPRIQVEHTVT 310
Db 241 EAPSPIMTEELRERMGAAIKAGKAINYDSAGTVEFLY-ENGNYFYLEMNTRIQVEHTVT 299
QY 311 BEVTEVDLVKAQMLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNGFRPDTGTTA 370
Db 300 EQVTGIDLVKAMIKIAAG--EELTLKQEDVKIRGHAIECRINAEDPLNDFVPCGKIKL 356
QY 371 YRSPGGAGVRLDGAALGGEITAHFDSMLVKMTCRSGDSDFETAVARAQALAEFTVSGVAT 430
Db 357 YRSPGGGVRIDSGVVGAEIPPYDYSMIKLITYGNSREEAIAEMKRALREYVIIGVKT 416
QY 431 NICFLRALLREEDFTSKRIATGFIADHPHLIQ 462
Db 417 NIPFHRVLEENFLKGNISTHYVEQNMHKLRL 448

RESULT 8
ACCC_ANASP STANDARD; PRT; 447 AA.
AC Q06862;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Biotin carboxylase (EC 6.3.4.14) (A subunit of acetyl-CoA
DE carboxylase) (EC 6.4.1.2) (ACC).
GN ACCC OR ALR0939.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93352435; PubMed=8102363;
RA Gornicki P., Scappino L.A., Haselkorn R.;
RT "Genes for two subunits of acetyl coenzyme A carboxylase of Anabaena
RT sp. strain PCC 7120: biotin carboxylase and biotin carboxyl carrier
RT protein.";
RL J. Bacteriol. 175:5268-5272(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
```

```
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: This protein is a component of the acetyl coenzyme A
CC carboxylase complex; first, biotin carboxylase catalyzes the
CC carboxylation of the carrier protein and then the transcarboxylase
CC transfers the carboxyl group to form malonyl-CoA.
CC -!- CATALYTIC ACTIVITY: ATP + biotin-carboxyl-carrier protein + CO(2)
CC = ADP + phosphate + carboxybiotin-carboxyl-carrier protein.
CC -!- PATHWAY: Long-chain fatty acid biosynthesis; first step.
CC -!- SUBUNIT: ACETYL-COA CARBOXYLASE IS AN HETEROHEXAMER OF BIOTIN
CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE AND THE TWO SUBUNITS
CC OF CARBOXYL TRANSFERASE IN A 2:2 COMPLEX.
CC -!- SIMILARITY: TO OTHER BIOTIN-DEPENDENT ENZYMES AND CARBAMOYL-
CC PHOSPHATE SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L14862; AAB51770.1; -.
CC EMBL; AP003584; BAB72896.1; -.
DR PIR; A53311; A53311.
DR PIR; AH1923; AH1923.
DR HSSP; P24182; 1BNC.
DR InterPro; IPR004549; AccC.
DR InterPro; IPR005482; Biotin carb_C.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005481; CPase_L_N.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR TIGRFAMs; TIGR00514; accC; 1.
DR PROSITE; PS00866; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
KW Fatty acid biosynthesis; Ligase; Biotin; ATP-binding;
KW Complete proteome.
FT NP_BIND 163 168 ATP (BY SIMILARITY).
FT ACT_SITE 293 293 BY SIMILARITY.
SQ SEQUENCE 447 AA; 49104 MW; 8A541B38B39E00F9 CRC64;

Query Match 17.9%; Score 1036.5; DB 1; Length 447;
Best Local Similarity 48.4%; Pred. No. 2.1e-48;
Matches 216; Conservative 73; Mismatches 152; Indels 5; Gaps 3;

QY 12 FKKILVANRGEIAVRAFAALETGAATVAIYPREDRGSFHRSFASAVRIGTEGSPVKAY 71
Db 3 FDKILIANRGEIALRILRACEEMGIATIAVHSTVDNALHVLQADEAVCIG-EPASAKSY 61
QY 72 LDIDEIIGAAGKVKADAIYPGYGFLSENAQLARECAENGITFTGPTPEVLDLTGDKSRV 131
Db 62 LNIPIIIAAALTRNASAIHPGYGFLSENAKFAEICADHIIAFIGTPEAIRLMGDKSTAK 121
QY 132 TAAKAGLPVLAESTPSKNID-EIVKSAEQTYPIFVKAVAGGGGRGMRFVASPDRLKL 190
Db 122 ETMQKAGVPTVPGSEGLVETEQEGLELAKDIGYPMIKATAGGGGRMLVRSDFEVLK 181
QY 191 ATEASREAEAAFGDGAIVYVERAVINPQHIEVQILGDHTGEVTVHLYERDCSLQRHOKVVE 250
Db 182 FLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQILADNYGNVHLDGDCSIQRRNQKLE 241
QY 251 IAPAQHLDPEDLRICADAVKFCRSIGYQAGTVEFLVDEKGNHVFIEHNPRIQVEHTVT 310
Db 242 EAPSPALDSDLREKMGQAQVAAQAFINYTGTACTIEFLLDRLSGQFYFMEHNTRIQVEHPT 301
```


